MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AT URBANA-CHAMPAIGN

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A124 494	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* ADJUSTMENT OF DESIGN PARAMETERS FOR IMPROVED FEEDBACK PROPERTIES IN THE LINEAR QUADRATIC REGULATOR.		5. TYPE OF REPORT & PERIOD COVERED Technical Report;
		6. PERFORMING ORG. REPORT NUMBER R-941(DC-50);UILU-ENG-2207
7. AUTHOR(*s*) ROBERT STANLEY MCEWEN		8. CONTRACT OR GRANT NUMBER(*s*) N00014-79-C-0424
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program		12. REPORT DATE March 1982
		13. NUMBER OF PAGES 91
14. MONITORING AGENCY NAME & ADDRESS(*if different from Controlling Office*)		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Linear quadratic
Robust
Multivariable control
Singular value

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Classical analysis and design methods for single input-single output (SISO) systems, such as gain and phase margins, do not generalize easily to MIMO systems. Recently, the singular values of the return difference and inverse Nyquist matrices have proven useful in analyzing multiple input-multiple output (MIMO) systems. The linear quadratic formulation is useful for the design of MIMO controllers. A disadvantage of this design method is that all the design specifications must be incorporated into a quadratic cost functional. This

DD FORM 1473 1 JAN 73   EDITION OF 1 NOV 65 IS OBSOLETE

ABSTRACT (continued)

thesis contains a systematic method for adjusting the quadratic cost to manipulate the singular value functionals and the feedback properties and thus achieve the design requirements.

UILU-ENG-82-2207

ADJUSTMENT OF DESIGN PARAMETERS FOR IMPROVED FEEDBACK
PROPERTIES IN THE LINEAR QUADRATIC REGULATOR

by

Robert Stanley McEwen

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special

A

ADJUSTMENT OF DESIGN PARAMETERS FOR IMPROVED
FEEDBACK PROPERTIES IN THE LINEAR QUADRATIC REGULATOR

BY

ROBERT STANLEY MCEWEN

B.S., Purdue University, 1978

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1982

Thesis Adviser: Professor Douglas P. Looze

Urbana, Illinois

iii

TABLE OF CONTENTS

1. INTRODUCTION

A major problem in the application of control theory is the fact that any mathematical model used for design or analysis is only an approximation to the true physical system. The error between the mathematical model and the system itself (or plant) has many causes. For example, coefficients that are assumed constant in the model may in reality be time varying. Also, the system may be too complex to be accurately described by a mathematical model that is feasible for design and analysis purposes. This is usually the case since a feasible model is typically restricted to be linear, time invariant, and finite dimensional.

Consequently, it is important to be able to design a controller so that the plant model-controller system (or nominal system) is as tolerant to plant variations as possible. This tolerance is important in two ways. First, the nominal system must remain stable under the expected range of plant variations. This system quality is often referred to as the robustness of the system. Second, the input output (I/O) characteristics of the nominal system should be as insensitive as possible to these variations. The nominal system is said to have good sensitivity properties if the effect of these variations on the I/O response is reduced from the effect on the plant alone.

The field of control theory is based upon the fact that feedback can be used to improve these and other important system qualities [1],[2]. For example, the effect of uncontrollable plant disturbances on the output of the plant can be reduced by feedback. Feedback controllers can also be

used to achieve a desirable I/O response, although this in itself does not necessarily require feedback [3].

The synthesis of closed loop controllers for single-input single-output (SISO) systems with these properties is well understood. Root locus plots, Nyquist diagrams, and other classical frequency domain techniques are readily used for design. The Nyquist diagram is particularly helpful since the distance from the Nyquist locus to the critical point (called the return difference function) provides a measure of the disturbance rejection and command following properties as a function of frequency. In addition, when the plant uncertainties are modeled by additive perturbations the return difference function is also a measure of robustness and sensitivity. When the uncertainties are modeled by multiplicative perturbations, the distance of the inverse Nyquist locus from the critical point provides this measure.

Unfortunately, these well-tested design techniques for SISO systems do not have an easy generalization to multiple-input multiple-output (MIMO) systems. Several problems unique to MIMO systems arise. For example, manipulations with transfer function matrices are more difficult since matrix multiplication is not commutative. Thus robustness and sensitivity margins are dependent on where the perturbation is inserted in the loop. Also, the perturbations not only have a fequency dependence, but a spatial dependence as well. In other words, uncertainties may occur only in certain loops.

Despite these problems, some useful generalizations of the classical theory to MIMO systems have recently been made [4-7]. These generalizations involve the return difference matrix or the inverse Nyquist matrix (these are the MIMO counterparts of the scalar functions). For example, when the return difference matrix of a nominally stable,

additively perturbed system is singular at some frequency, it means that a zero pair of the characterisitc equation lies on the imaginary axis. Thus for all stable additive perturbations that are smaller than the smallest perturbation that causes singularlity, the system will be stable. Hence the distance of the return difference matrix from singularity provides a measure of the stability margin of the system under stable additive perturbations. This is a direct generalization of the SISO case where the distance of the Nyquist locus from the critical point represents the degree of stability, which is measured by the classical gain and phase margins. Furthermore, it can be shown that the return difference and inverse Nyquist matrices also reflect the sensitivity, disturbance rejection, and I/O properties of MIMO systems.

When the standard Euclidean coordinates and corresponding norm are used to describe the system, this distance from singularity can be conveniently computed in terms of matrix singular values. Several singular value inequalities for robustness and sensitivity under additive or multiplicative perturbations are available [8].

The design requirements on robustness, sensitivity, disturbance rejection and I/O response translate into bounds on these singular value curves (as functions of frequency). Thus a major design objective is to synthesize a controller that adjusts the singular value curves of the return difference or inverse Nyquist matrix into some desirable shape. If the mathematical structure of the controller is predetermined, it is possible to calculate the gradient of these singular values with respect to the adjustable parameters of the controller (or control parameter vector). The gradient will be a function of the control parameter vector and frequency.

Given an initial control parameter vector and a constant frequency, the
gradient indicates how to change the initial control parameter vector to
have the greatest possible effect on the singular value curve at that
frequency.

In order to deform the entire singular value curve to fit specifi-
cations, it will generally be necessary to follow an iterative select-
evaluate-adjust design procedure. Each time a new control parameter vector
is selected, the singular values must be analyzed. If the singular values
are not satisfactory, the gradient can be recalculated at the frequency of
interest, and the control parameter vector readjusted. Then the singular
values are analyzed again, and so on.

The success of the design depends on the capability of the designer.
to use the gradient information along with his insight into the particular
system to determine how to adjust the parameter vector to set the desired
effect on the closed loop system. This design method is obviously highly
dependent on the particular problem under consideration. However, if the
designer is using a standardized synthesis method at each design iteration
to adjust the controller, such as the linear quadratic (LQ) design method,
the gradient computations can be formulated explicitly.

The solution of the full state feedback LQ problem is a constant
gain feedback matrix that optimizes a cost functional which contains weighted
quadratic state and control terms [9]. These weights are the choice of the
designer, and determine the feedback gain matrix. An advantage of the LQ
design method is that it provides a systematic, numerically feasible method
for choosing a feedback gain matrix with relatively small gains (if all of
the state weights are set to zero then the minimum energy regulator solution

results). By scaling these weights the LQ method can be used for approximate pole placement [10]. Also, the LQ loop has inherently good robustness and sensitivity properties [11],[12].

A disadvantage of the LQ design is that it is not directly obvious how to choose these weights to achieve a given set of design specifications. In particular, it is not at all obvious how to choose these weights to alter the singular values of the return difference and inverse Nyquist matrices, and thus affect the robustness and sensitivity properties of the system.

In this thesis, explicit formulas for the gradient of the singular value functionals of the return difference and inverse Nyquist matrices are derived for the LQ problem. This gradient information can be used iteratively to tune the state and control weights to shape the singular values as a function of frequency and thus obtain desired robustness and sensitivity properties.

In Section 2, basic properties of singular values are reviewed. then the applications of singular value analysis in control theory such as multivariable Bode plots and robustness-sensitivity bounds are discussed. The Lyapunov operator is defined, and the differential and gradient concepts are reviewed. Sensitivity formulas for the eigenvalues and singular values of a matrix are developed.

Section 3 contains a discussion on MIMO loop shaping in the frequency domain, and how desirable system properties translate into requirements on the smallest singular value of the return difference matrix and the inverse Nyquist matrix. Following this, general sensitivity formulas for these singular value functionals are derived, and an iterative design method

is proposed for manipulating these functionals. This theory is then applied to the LQ problem to obtain specific results.

Finally, Section 4 presents an example application of the method developed in Section 3. The design method of Section 3 and an asymptotic design method are compared. Section 5 summarizes the thesis and discusses further research possibilities.

2. SINGULAR VALUES IN CONTROL THEORY

In this section the basic background needed for the presentation of Section 3 is reviewed. Singular values and their role in the analysis of robustness and sensitivity of feedback systems are discussed in general. Vector space and gradient concepts are reviewed as a preliminary to differential eigenvalue and singular value formulas.

2.1. The Singular Value Decomposition and Near Singular Matrices

In control theory it is often desirable to measure the nearness to singularity of a matrix. The following theorem provides such a measure.

Theorem 2.1: Suppose that A and ΔA are n×n matrices, and that the inverse A^{-1} exists. Then, the inverse

$$(A + \Delta A)^{-1} \tag{2.1}$$

exists if

$$\| \Delta A \| < \| A^{-1} \|^{-1} \tag{2.2}$$

where $\| \cdot \|$ is the standard induced Euclidean norm. □

Proof: [13].

The matrix ΔA is an arbitrary additive perturbation. Since the inverse in equation (2.1) will exist for every perturbation that satisfies equation (2.2), the magnitude of the functional $\| A^{-1} \|^{-1}$ indicates how sensitive a matrix is to changes in its entries. For example, if the value of $\| A^{-1} \|^{-1}$ is small, then a small change in an element of A could cause singularity.

Notice also that equation (2.2) is only a sufficient condition. This means that the matrix $A + \Delta A$ may or may not be singular for a given perturbation that violates equation (2.2).

A singular value of an arbitrary complex valued m×n matrix A is denoted by $\sigma_i(A)$, and is defined by

$$\sigma_i(A) \triangleq \sqrt{\overline{\lambda_i(A^H A)}} \qquad (2.3)$$

for each

$$i = 1, \ldots, \ell$$

where

$$\ell \triangleq \min\{m, n\}.$$

Here, $\lambda_i(\cdot)$ is the ith eigenvalue of the indicated matrix. Notice that singular values are defined for nonsquare matrices in general. Also, singular values are always nonnegative real numbers since any positive semidefinite hermitian matrix has eigenvalues that are real and nonnegative.

<u>Theorem 2.2</u>: Any matrix $A \in C^{m \times n}$ of rank ℓ has a singular value decomposition given by

$$A = U\Sigma V^H \qquad (2.4)$$

where $U \in C^{m \times \ell}$ and $V \in C^{n \times \ell}$ are unitary matrices and where

$$\Sigma = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_\ell\} \qquad (2.5)$$

with

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_\ell. \qquad (2.6)$$

<u>Proof</u>: [13]. □

From equation (2.4),

$$A^H A V = V\Sigma^2. \qquad (2.7)$$

Hence the columns of V are orthonormal eigenvectors of the hermitian matrix $A^H A$. These vectors are denoted v_i and are called the right singular vectors. By convention, the v_i in equation (2.7) are ordered so that inequality (2.6) is satisfied. Also from equation (2.4),

$$AA^H U = U\Sigma^2. \tag{2.8}$$

The columns of U are denoted u_i and are termed the left singular vectors. Some useful properties of singular vectors are given below.

Corollary 2.3:

(1) The right and left singular vectors are related by

$$\begin{aligned} Av_i &= \sigma_i u_i \\ A^H u_i &= \sigma_i v_i. \end{aligned} \tag{2.9}$$

(2) If A is invertible, where $A \in \mathbb{C}^{n \times n}$, then the singular values of A^{-1} are related to those of A by

$$\sigma_i(A^{-1}) = \frac{1}{\sigma_{n+1-i}(A)} \quad , \quad i = 1, \ldots, n. \tag{2.10}$$

(3) The induced Euclidean norm is given by

$$\|A\|_2 = \sigma_1(A). \tag{2.11}$$

Proof: [13]. □

Property (2) allows Theorem 2.1 to be interpreted in terms of singular values. Equation (2.2) is equivalent to

$$\bar{\sigma}(\Delta A) < \underline{\sigma}(A^{-1}) \tag{2.12}$$

where

$$\begin{aligned} \bar{\sigma}(\cdot) &\triangleq \sigma_1(\cdot) \\ \underline{\sigma}(\cdot) &\triangleq \sigma_n(\cdot). \end{aligned} \tag{2.13}$$

Inequality (2.12) plus the fact that there exists a perturbation ΔA such that $\bar{\sigma}(\Delta A) = \underline{\sigma}(A)$ and $A + \Delta A$ is singular means that the smallest singular value of a matrix is the distance from that matrix to the nearest singular matrix.

2.2. Singular Values in Control Theory

Suppose that $P^*(j\omega)$ is a m×n rational transfer function matrix of a multiple input-multiple output (MIMO) linear system, with

$$y(j\omega) = P(j\omega)u(j\omega), \quad y \in \mathbb{C}^n, \quad u \in \mathbb{C}^m. \tag{2.14}$$

Then, it has been shown [14] that for all ω,

$$\underline{\sigma}(P^*(j\omega)) \leq \frac{\|y(j\omega)\|}{\|u(j\omega)\|} \leq \bar{\sigma}(P^*(j\omega)) \tag{2.15}$$

where $\bar{\sigma}(P^*(j\omega))$ is the least upper bound and $\underline{\sigma}(P^*(j\omega))$ is the greatest lower bound of the above vector ratio.

If we let $\|u(j\omega)\| = 1$, then the output $\|y(j\omega)\|$ is bounded between the two singular values at a given ω for any input direction. A frequency plot of these singular values can be regarded as the *MIMO generalization of a single input-single output (SISO) Bode magnitude plot* (see Figure 2.1).

The interpretation of a Bode magnitude plot as a rms sinusoidal power gain curve also carrys over to the MIMO case [15]. Suppose that an input

$$u(t) = u \cos \omega t$$

(where u is a vector of constants) is applied to the system of equation (2.14). The sum of the mean squared power of each input signal is

$$SMSPI = \frac{\omega}{2\pi} \int_0^{\frac{2\pi}{\omega}} u^T(t)u(t)dt = \frac{1}{2} u^T u.$$

The sum of the mean squared power of the outputs

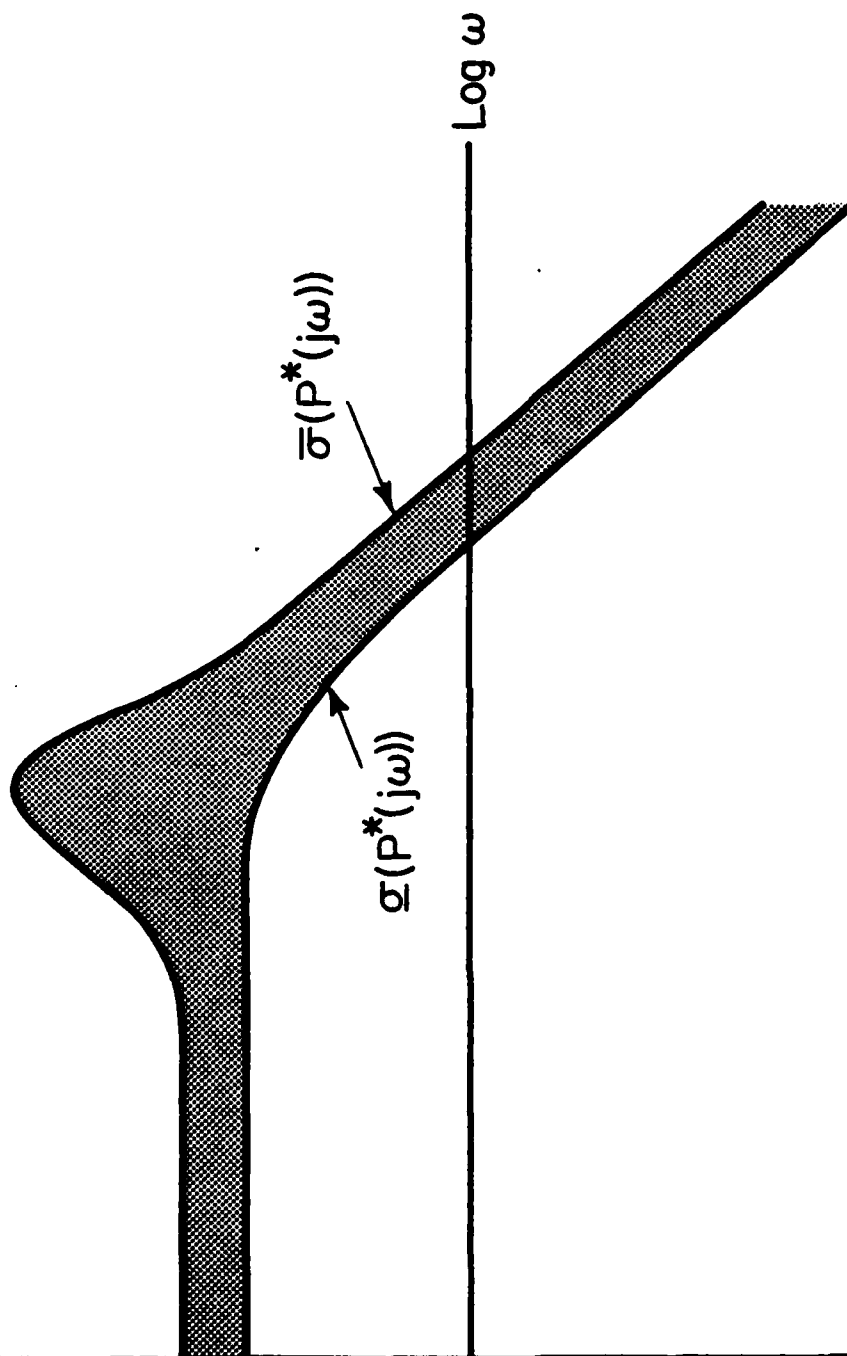$$SMSPO = \frac{\omega}{2\pi} \int_0^{\frac{2\pi}{\omega}} y_{ss}^T(t)y_{ss}(t)dt$$

Figure 2.1. Singular value magnitude plot for a MIMO system.

FP-7389

where $y_{ss}(t)$ is a vector of steady state (sinusoidal) outputs. Then, the ratio of mean squared power out to mean squared power in is bounded by the squares of the singular values of $P^*(j\omega)$ for every input direction u

$$\underline{\sigma}(P^*(j\omega))^2 \leq \frac{SMSPO}{SMSPI} \leq \bar{\sigma}(P^*(j\omega))^2.$$

Singular values have another extremely important interpretation in the robustness analysis of MIMO systems. A closed loop feedback system is said to be robust if it remains stable when the true plant, $P(j\omega)$ varys from the nominal model $P^*(j\omega)$ that was used to design the feedback controller.

This difference between the real plant and the mathematical representation can be modeled in many ways. Three models which have proven useful for analysis are:

(1) Additive perturbations

$$P(j\omega) = P^*(j\omega) + \Delta P_A(j\omega) \qquad (2.16)$$

where $P(j\omega)$ is the true plant, $P^*(j\omega)$ is the nominal model, and $\Delta P_A(j\omega)$ is an unknown quantity.

(2) Multiplicative input perturbations

$$P(j\omega) = P^*(j\omega)(I + \Delta P_M(j\omega)). \qquad (2.17)$$

(3) Multiplicative output perturbations

$$P(j\omega) = (I + \Delta P_M(j\omega))P^*(j\omega). \qquad (2.18)$$

Bounds expressed in terms of singular values have been derived that relate the nominal plant to the magnitude of the perturbation that can be tolerated before stability is no longer assured. Consider the case of additive perturbations (see Figure 2.2) where $P^*(j\omega)$ is an m×n nominal plant transfer function, and $K(j\omega)$ is the n×m controller.
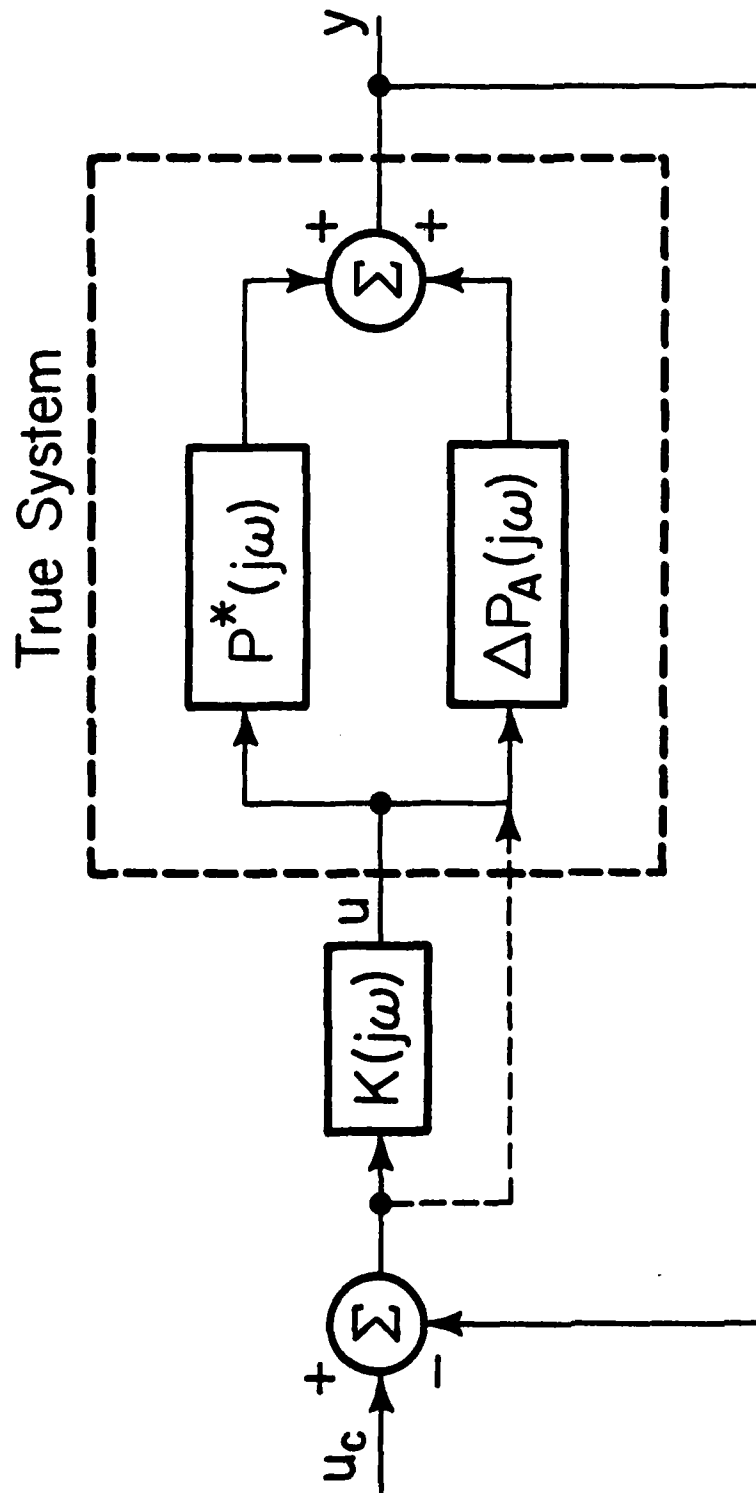
Figure 2.2. Additively perturbed system.

Theorem 2.4:   Assume :

(1) $P^*(j\omega)$, $\Delta P_A(j\omega)$, and $K(j\omega)$ are rational transfer function matrices

(2) $\Delta P_A(j\omega)$ is stable.

(3) The nominal closed loop system ($\Delta P_A(j\omega) \equiv 0$) is stable.

Then, the closed loop system described by solid lines in Figure 2 is stable for all $\Delta P_A(j\omega)$ which satisfy

$$\underline{\sigma}(I + P^*(j\omega)K(j\omega)) > \bar{\sigma}(\Delta P_A(j\omega))\bar{\sigma}(K(j\omega)) \qquad (2.19)$$

for all $\omega \in R$.

Proof: [15].   □

        The utility of this theorem lies in the fact that only knowledge of the norm of the perturbation matrix is required to ensure stability.

        Equation (2.19) can also be interpreted geometrically as a generalization of the Nyquist criterion for SISO systems. The left hand side of equation (2.19) is analogous to the distance of the SISO Nyquist locus from the critical point -1+j0 in the complex plane [15].

        A similar theorem holds in the case of multiplicative perturbations. Consider the system in Figure 2.3.

Theorem 2.5: Assume :

(1) $P^*(j\omega)$, $\Delta P_M(j\omega)$, and $K(j\omega)$ are rational transfer function matrices.

(2) $\Delta P_M(j\omega)$ is stable.

(3) The nominal closed loop system ($\Delta P_M(j\omega) \equiv 0$) is stable.

(4) $\det\{P^*(j\omega)K(j\omega)\} \neq 0$.

Then the perturbed system is stable for all $P_M(j\omega)$ which satisfy

$$\underline{\sigma}(I + [P^*(j\omega)K(j\omega)]^{-1}) > \bar{\sigma}(\Delta P_M(j\omega)) \qquad (2.20)$$

for all $\omega \in R$.

Proof: [4].   □

True System

$$u_c \xrightarrow{+} \Sigma \xrightarrow{} K(j\omega) \xrightarrow{u} P^*(j\omega) \xrightarrow{} I + \Delta P_M(j\omega) \xrightarrow{} y$$
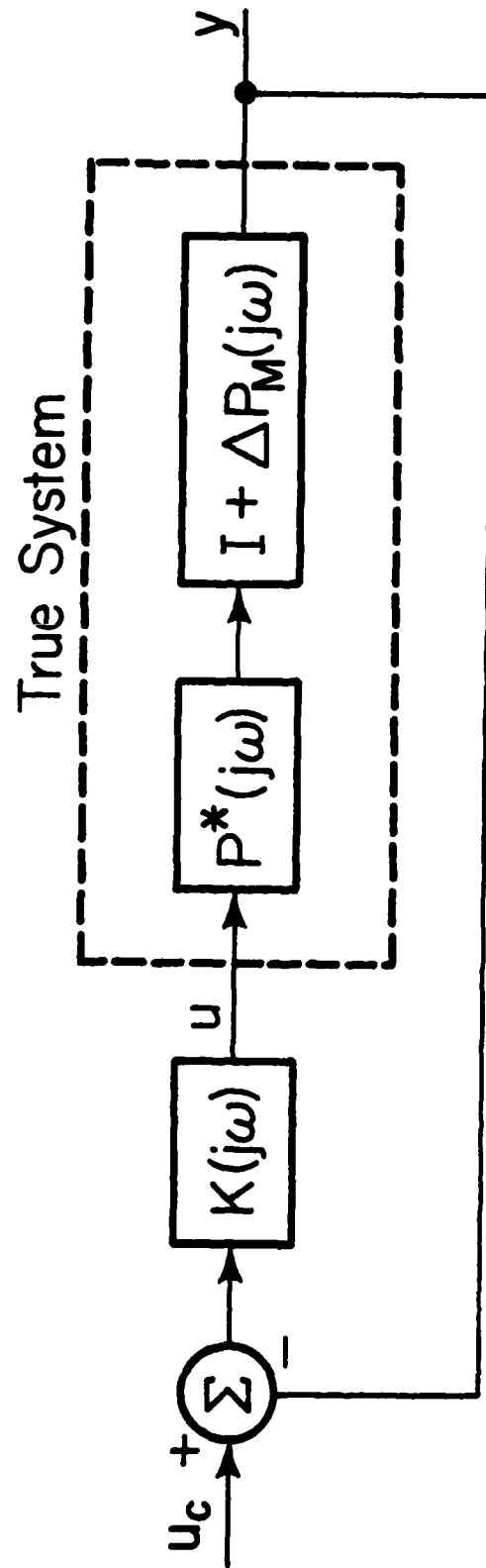
FP-7382

Figure 2.3. Multiplicatively perturbed system.

Here, the robustness is measured with respect to the inverse Nyquist matrix $I + [P^*(j\omega)K(j\omega)]^{-1}$ rather than the return difference matrix as in equation (2.19).

It is important to be aware that equation (2.20) is valid only for perturbations that are at the output of the plant, as shown in Figure 2.3. If the perturbation was applied at the input of the plant (the positions of $P^*(j\omega)$ and $I + \Delta P_M(j\omega)$ would be reversed in Figure 2.3) equation (2.20) becomes

$$\underline{\sigma}(I + [K(j\omega)P^*(j\omega)]^{-1}) > \bar{\sigma}(\Delta P_M(j\omega))$$

where the matrix $K(j\omega)P^*(j\omega)$ is n×n rather than m×m. For the sake of consistency, all of the following theorems in this section will deal with output perturbations. Results for input perturbations are similar and can be derived from the references given for each theorem.

The above theorem shows how multiplicative perturbations impose limits on the size of the closed loop bandwidth ω_B. (We define the bandwidth of a MIMO system as the frequency at which the largest singular value of the transfer function matrix drops to $1/\sqrt{2}$ of its zero frequency value.) The relation

$$[I + [P^*(j\omega)K(j\omega)]^{-1}]^{-1} = P^*(j\omega)K(j\omega)[I + P^*(j\omega)K(j\omega)]^{-1} \qquad (2.21)$$

plus equation (2.10) implies that the smallest singular value of the inverse Nyquist matrix is just the reciprocal of the largest singular value of the transfer function matrix. At frequencies where large multiplicative perturbations are possible, the above theorem requires that

$$\underline{\sigma}(I + (P^*(j\omega)K(j\omega)]^{-1})$$

be large, and hence

$$\bar{\sigma}(P^*(j\omega)K(j\omega)[I + P^*(j\omega)K(j\omega)]^{-1})$$

be small.

The effect of modeling uncertainties and parameter variations on input-output response in MIMO system can also be measured in terms of singular values. One way to do this that provides a logical generalization of the SISO case is to define a nominally equivalent open loop system, and then deduce conditions that require the output of the closed loop system to be less sensitive to perturbations [5]. Consider the system described by solid lines in Figure 2.4 where

$$K_o(j\omega) \stackrel{\Delta}{=} K(j\omega)(I + P^*(j\omega)K(j\omega))^{-1} . \qquad (2.22)$$

Comparing this open loop system to the closed loop system of Figure 2.2, we see that when $\Delta P_A(j\omega) \equiv 0$ the two systems will respond identically for any given input U_c. For the case $\Delta P_A(j\omega) \equiv 0$, $y_o(j\omega)$ will denote the output of the open loop system, and y_c will denote the output of the closed loop system. When $\Delta P_A(j\omega) \neq 0$, these outputs will be denoted by $y_o'(j\omega)$ and $y_c'(j\omega)$. The open loop and closed loop errors are defined by

$$e_o(j\omega) = y_o(j\omega) - y_o'(j\omega)$$

$$e_c(j\omega) = y_c(j\omega) - y_c'(j\omega). \qquad (2.23)$$

The open loop and closed loop systems can be compared through the mean squared errors, which are defined to be
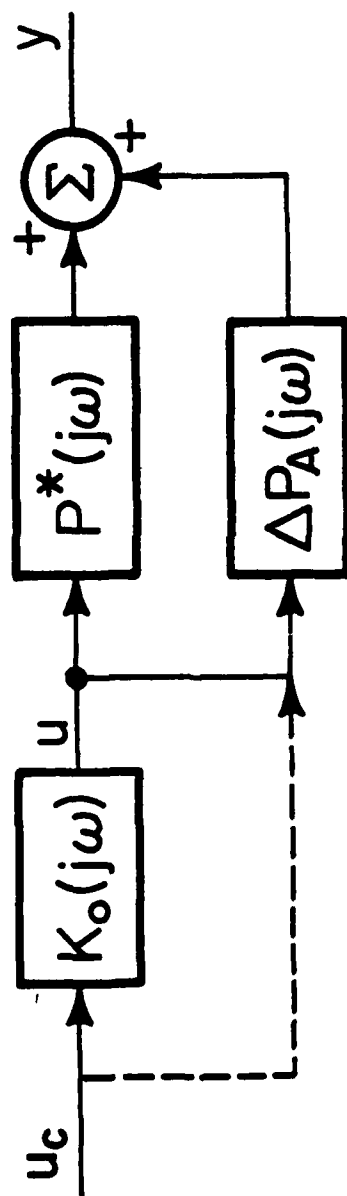
FP-7383

Figure 2.4. Nominally equivalent additively perturbed open loop system.

$$J_o = \int_0^\infty \|e_o(t)\|^2 dt$$

$$J_c = \int_0^\infty \|e_c(t)\|^2 dt \qquad (2.24)$$

where $e_o(t)$ and $e_c(t)$ are the inverse Fourier transforms of equation (2.23).

Theorem 2.6: Suppose:

(1) $K(j\omega)$, $P^*(j\omega)$, and $\Delta P_A(j\omega)$ are rational transfer function matrices.

(2) $\Delta P_A(j\omega)$ and $(I + P^*(j\omega)K(j\omega))^{-1}$ are stable.

(3) The perturbed system of Figure 2.2 is stable (i.e. $\Delta P_A(j\omega)$ satisfies equation (2.19)).

If $\Delta P_A(j\omega)$ satisfies

$$\underline{\sigma}(I + P^*(j\omega)K(j\omega)) \geq \bar{\sigma}(\Delta P_A(j\omega))\bar{\sigma}(K(j\omega)) + 1 \qquad (2.25)$$

for all ω in some interval $\mathcal{J}$, and $u_c(j\omega) = 0$ for all $\omega \notin \mathcal{J}$, then

$$J_c \leq J_o. \qquad (2.26)$$

Proof: [8].

When the conditions of this theorem are satisfied, the output of the closed loop system in Figure 2.2 will be less sensitive to additive perturbations than the output of the equivalent open loop system. Under these conditions, feedback has a desensitizing effect.

The sensitivity condition (2.25) is remarkably similar to the stability condition (2.20). Indeed, if a system has good sensitivity properties under additive perturbations, it must necessarily also have good stability margins under additive perturbations in the frequency band $\mathcal{J}$.

A similar result holds for multiplicative perturbations. Consider the multiplicatively perturbed systems in Figures 2.3 and 2.5, with errors defined as in equation (2.23).
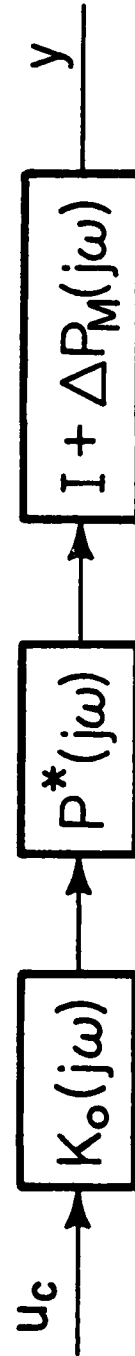
FP-7384

Figure 2.5. Nominally equivalent multiplicatively perturbed open loop system.

Theorem 2.7: Suppose

(1) $\Delta P_M(j\omega)$, $P^*(j\omega)$, and $K(j\omega)$ are rational transfer function matrices.

(2) $\Delta P_M(j\omega)$ and $(I + P^*(j\omega)K(j\omega))^{-1}$ are stable.

(3) $\Delta P_M(j\omega)$ satisfies equation (2.20).

(4) $\det\{P^*(j\omega)K(j\omega)\} \not\equiv 0$.

If $\Delta P_M(j\omega)$ satisfies

$$\underline{\sigma}(I + [P^*(j\omega)K(j\omega)]^{-1}) \geq \bar{\sigma}(\Delta P_M(j\omega)) + \bar{\sigma}([P^*(j\omega)K(j\omega)]^{-1}) \qquad (2.27)$$

for all ω in some interval $\mathcal{J}$, and $u_c(j\omega) = 0$ for all $\omega \notin \mathcal{J}$, then

$$J_c \leq J_o.$$

Proof: [8]. □

Since $\bar{\sigma}([P^*(j\omega)K(j\omega)]^{-1})$ is a nonnegative number, we again see that good sensitivity properties of a multiplicatively perturbes system imply good stability margins (see equation (2.20)) for $\omega \in \mathcal{J}$.

Also, the sensitivity conditions (2.25) and (2.27) need only be satisfied in a finite bandwidth. The robustness conditions (2.19) and (2.20) must be satisfied for all ω.

Another important fact is the the robustness sensitivity theorems for additive perturbations (Theorems 2.4 and 2.6) are conservative. For example, there does not necessarily exist a destabilizing $\Delta P_A(j\omega)$ which will satisfy equation (2.19) with equality. However, if additive perturbations are considered in the configuration indicated by the dotted line in Figures 2.2 and 2.4, the bounds (2.19) and (2.20) are not conservative.

2.3. Linear Operators and the Lyapunov Equation

It is easily demonstrated that the collection of all m n complex valued matrices ($\mathbb{C}^{m \times n}$) is a linear vector space. A useful inner product on this space is

$$\langle M_1, M_2 \rangle = \mathrm{tr}\{M_1^H, M_2\} \quad , \qquad M_1, M_2 \in \mathbb{C}^{m \times n}. \tag{2.28}$$

The natural norm associated with this inner product is the Frobenius norm

$$\begin{aligned}
\sqrt{\langle M_1, M_1 \rangle} &= \sqrt{\mathrm{tr}\{M_1^T M_1\}} \\
&= \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |m_{ij}|^2} \\
&\overset{\Delta}{=} \|M_1\|_F.
\end{aligned} \tag{2.29}$$

A vector space is complete under a given norm if every Cauchy sequence of vectors in the space converges to a vector which is also a member of the space. A complete, normed, linear vector space is a Banach space. If, in addition, an inner product which induces the norm is defined, the space is a Hilbert space. Under the inner product defined in equation (2.28). $\mathbb{C}^{m \times n}$ is a Hilbert space.

A concept that will be useful later is cartesian product Hilbert space. A cartesian product of two Hilbert spaces is a collection of all ordered pairs consisting of one vector from the first space and one vector from the second. For example,

$$S \overset{\Delta}{=} \mathbb{C}^{m \times n} \times \mathbb{C}^{m \times m} \tag{2.30}$$

is a cartesian product space. An element of this space is written

$$(M,N), \quad M \in \mathbb{C}^{m \times n}, \quad N \in \mathbb{C}^{m \times m}.$$

An inner product is defined on a Cartesian product space by adding the inner products from each constituent space. For example, the inner product on S is

$$\langle (M_1,N_1),(M_2,N_2) \rangle = \langle M_1,M_2 \rangle + \langle N_1,N_2 \rangle$$
$$= tr\{M_1^H M_2\} + tr\{N_1^H N_2\}. \qquad (2.31)$$

Under this inner product the Cartesian product of two Hilbert spaces is itself a Hilbert space.

The concepts of operators and adjoints will be useful in the derivation of the gradient in the following section. Suppose $\mathcal{J}:\mathcal{X} \to \mathcal{Y}$ is a linear operator where $\mathcal{X}$ and $\mathcal{Y}$ are inner product spaces. If there exists another operator $\mathcal{J}^*:\mathcal{Y} \to \mathcal{X}$ that satisfies

$$\langle y,\mathcal{J}(x) \rangle = \langle \mathcal{J}^*(y),x \rangle \quad , \qquad \forall x \in \mathcal{X}, \; y \in \mathcal{Y}. \qquad (2.32)$$

then $\mathcal{J}^*$ is called the adjoint of $\mathcal{J}$. When $\mathcal{X}$ and $\mathcal{Y}$ are Hilbert spaces, the adjoint always exists and is unique. The adjoint operator has the following properties [16]:

(1) The adjoint operator $\mathcal{J}^*$ is linear.

(2) If $\mathcal{J}$ has an inverse $\mathcal{J}^{-1}$ then

$$(\mathcal{J}^{-1})^* = (\mathcal{J}^*)^{-1}. \qquad (2.33)$$

It will be of interest to us to examine the Lyapunov equation in the context of operators and adjoints. First, recall the following fundamental properties of the Lyapunov equation.

<u>Theorem 2.10</u>: Consider the Lyapunov equation

$$KA + A^H K + Q = 0 \qquad (2.34)$$

where

$$A, K, Q \in \mathbb{C}^{n \times n}$$

and A is stable (i.e. $\text{Re}\{\lambda_i(A)\} > 0$, $i=1,\ldots,n$). Then

(1)   For any $Q \in \mathbb{C}^{n \times n}$ there exists a unique solution $K \in \mathbb{C}^{n \times n}$.

(2)   If Q is Hermitian then K is Hermitian.

(3)   If Q is positive semidefinite then K is positive semidefinite.

Proof:

(1)   [17]

(2)   This follows directly from (1).

(3)   [17].   □

Since the Lyapunov equation assigns a unique solution K to every input Q, it can be thought of as an operator that maps from $\mathbb{C}^{n \times n}$ to $\mathbb{C}^{n \times n}$. Define the linear operator L_A as

$$L_A(K) \overset{\Delta}{=} A^H K + AK \qquad (2.35)$$

where A is always assumed to be stable. Equation (2.34) is then equivalent to

$$L_A(K) = -Q. \qquad (2.36)$$

This linear operator has several very useful properties. First, it follows from Theorem 2.10 that $L_A(\cdot)$ is one to one and onto. Therefore an inverse Lyapunov operator exists and is linear. It will be denoted $L_A(\cdot)^{-1}$, where

$$L_A^{-1}(Q) = -K. \qquad (2.37)$$

Second, since $\mathbb{C}^{n \times n}$ is a Hilbert space under the inner product defined in equation (2.28), the adjoint operator of $L_A(\cdot)$ exists.

<u>Theorem 2.12</u>: Consider the linear operator L_A on the Hilbert space $\mathbb{C}^{n \times n}$ with the inner product defined in equation (2.28). Then, the adjoint operator $L_A^*(\cdot)$ is given by

$$L_A(\cdot)^* = L_A^H(\cdot). \tag{2.38}$$

<u>Proof</u>: From equations (2.28) and (2.32) we have

$$\langle Y, L_A(K) \rangle = \text{tr}\{Y^H L_A(K)\}, \qquad Y, K \in \mathbb{C}^{n \times n}.$$

By the definition of L_A,

$$\langle Y, L_A(K) \rangle = \text{tr}\{Y^H(KA + A^H K)\}$$
$$= \text{tr}\{Y^H KA\} + \text{tr}\{Y^H A^H K\}.$$

Using elementary trace properties,

$$\langle Y, L_A(K) \rangle = \text{tr}\{A^H K^H Y\} + \text{tr}\{K^H AY\}$$
$$= \text{tr}\{K^H YA^H\} + \text{tr}\{K^H AY\}$$
$$= \text{tr}\{K^H(YA^H + AY)\}$$
$$= \text{tr}\{K^H L_A^H(Y)\}$$
$$= \langle K, L_A^H(Y) \rangle$$

which implies $L_A(\cdot)^* = L_A^H(\cdot)$ by equation (2.32).   □

Another concept which will be useful in the next chapter is the square root of a matrix. A matrix M is a square root of a symmetric matrix Q whenever

$$M^T M = Q. \tag{2.39}$$

Under certain conditions, the square root M will always exist, although it is in general nonunique.

<u>Theorem 2.13</u>: Suppose that Q is a n×n real, symmetric, positive semi-definite matrix with rank m. Then there exists a m×n square root.

<u>Proof</u>: Straightforward. □

Observe that if Q has an m×n square root M, then there will be a ℓ×n square root for every ℓ>m.

Finally, the following result pertaining to differentiable matrices is necessary for the derivations in the next chapter.

<u>Theorem 2.14</u>: If V(t) is an n×n matrix whose entries are real valued, differentiable functions of t, and the inverse $V^{-1}(t)$ exists for all t, then

$$\dot{V}^{-1}(t) = -V^{-1}(t)\dot{V}(t)V^{-1}(t). \tag{2.40}$$

<u>Proof</u>: The proof is obvious from the fact that

$$V(t)V^{-1}(t) = I$$

and the product rule for differentiation. □

2.4. The Differential

The concept of the scalar differential can be extended to operators and vector spaces. Suppose that

$$\mathscr{F}: \mathscr{X} \to \mathscr{Y} \tag{2.41}$$

where $\mathscr{X}$ and $\mathscr{Y}$ are Hilbert spaces. If there is an operator $\delta\mathscr{F}(x;\Delta x)$ that is linear and continuous with respect to $\Delta x \in \mathscr{X}$, and if $\delta\mathscr{F}(x;\Delta x)$ satisfies

$$\lim_{\|\Delta x\| \to 0} \frac{\|\mathscr{F}(x+\Delta x) - \mathscr{F}(x) - \delta\mathscr{F}(x,\Delta x)\|}{\|\Delta x\|} = 0 \tag{2.42}$$

then $\delta\mathscr{F}(x;\Delta x)$ is the Frechet differential of $\mathscr{F}$ at x in the direction Δx.

Notice that the differential assigns a vector y to every ordered pair $(x, \Delta x)$,

$$\delta \mathcal{F} : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}. \tag{2.43}$$

Since $\delta \mathcal{F}(x; \Delta x)$ is linear in Δx, it can be written (for each $x \in \mathcal{X}$),

$$\delta \mathcal{F}(x; \Delta x) = \mathcal{F}'(x) \Delta x \tag{2.44}$$

where $\mathcal{F}'$ is a bounded linear operator that depends on x,

$$\mathcal{F}' : \mathcal{X} \to B(\mathcal{X}, \mathcal{Y}).$$

Here, $B(\mathcal{X}, \mathcal{Y})$ denotes the space of all bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$. $\mathcal{F}'$ is called the Fréchet derivative of $\mathcal{F}$ at x.

Many of the properties of scalar derivatives extend to Fréchet derivatives, such as the familiar chain rule.

<u>Theorem 2.16</u> (Chain Rule): Suppose that $\mathcal{F}$ and $\mathcal{H}$ are operators

$$\begin{aligned} \mathcal{F} &: \mathcal{X} \to \mathcal{Y} \\ \mathcal{H} &: \mathcal{Y} \to \mathcal{Z} \end{aligned} \tag{2.45}$$

where $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ are Hilbert spaces. Suppose further that the Fréchet derivative $\mathcal{F}'$ exists for all $x \in \mathcal{X}$ and that $\mathcal{H}'$ exists for all $y \in \mathcal{Y}$. Then, the operator

$$\mathcal{H} \cdot \mathcal{F} : \mathcal{X} \to \mathcal{Z} \tag{2.46}$$

has a Fréchet derivative given by

$$(\mathcal{H} \cdot \mathcal{F})'(x) = \mathcal{H}'(\mathcal{F}(x)) \, \mathcal{F}'(x). \tag{2.47}$$

<u>Proof</u>: [16]. □

The chain rule also applies to Fréchet differentials. It follows from equation (2.47) that the differential of the composite operator in equation (2.46) is

$$\delta \mathcal{K} \cdot \mathcal{A}(x; \Delta x) = \delta \mathcal{K}(\mathcal{A}(x); \delta \mathcal{A}(x; \Delta x)) \qquad (2.48)$$

where $\mathcal{K}$ and $\mathcal{A}$ are under the same assumptions as in the theorem.

The geometrical interpretation of the Fréchet differential is similar to that of the scalar differential. Equation (2.42) is equivalent to saying that for all $\epsilon > 0$, there exists $\delta(\epsilon) > 0$ such that if $\|\Delta x\| < \delta(\epsilon)$, then

$$\|\mathcal{A}(x+\Delta x) - \mathcal{A}(x) - \delta \mathcal{A}(x; \Delta x)\| < \epsilon \|\Delta x\| .$$

Since

$$\|\mathcal{A}(x+\Delta x) - \mathcal{A}(x)\| - \|\delta \mathcal{A}(x; \Delta x)\| \leq \|\mathcal{A}(x+\Delta x) - \mathcal{A}(x) - \delta \mathcal{A}(x; \Delta x)\|$$

we can see that for small $\|\Delta x\|$, the magnitude of the differential is approximately equal to the magnitude of the change of $\mathcal{A}(x)$ in the Δx direction. The differential, then, can be interpreted as a sensitivity function.

In the case of functionals, this concept can be carried further. First, recall that a functional is an operator that maps a vector space into the real line,

$$f : \mathcal{X} \to \mathbf{R}. \qquad (2.49)$$

Theorem 2.15 (Riesz Fréchet): If g is a bounded linear functional on a Hilbert space $\mathcal{X}$, then there exists a unique vector h in $\mathcal{X}$ such that for all $x \in \mathcal{X}$,

$$g(x) = \langle h, x \rangle$$

with

$$\|g\| = \|h\|$$

where $\|g\|$ is the induced functional norm,

$$\|g\| \overset{\Delta}{=} \sup_{\|x\|=1} \{|g(x)|\}.$$

<u>Proof</u>: [16].   □

The true power of this theorem is that it applies in infinite dimensional spaces. In this thesis we will be concerned primarily with finite dimensional Euclidean vector space. However, the above theorem is still useful because it provides a means of calculating the gradient through the differential.

Consider the functional f, as in equation (2.49), over a Hilbert space $\mathscr{H}$. Since the differential $\delta g(x;\Delta x)$ is linear in Δx, the Reisz-Fréchet theorem guarantees that there exists a vector $\nabla_x f$ such that

$$\delta f(x;\Delta x) = \langle \nabla_x f(x), \Delta x \rangle \qquad (2.50)$$

where $\nabla_x f(x_o)$ denotes the gradient of f with respect to x at the point x_o. Recall that the gradient is a vector that points in the direction of maximum increase of f at x_o, and whose magnitude is equal to the value of the directional derivative in that direction.

Equation (2.50) also provides another geometrical interpretation of the differential. When $\|\Delta x\| = 1$, the differential is just the gradient projected in an arbitrary Δx direction. In this case, the differential is equal to the directional derivative. This suggests a way to calculate the differential.

<u>Theorem 2.16</u>: If the Fréchet differential of the functional in equation (2.49) exists, then it is unique and is given by

$$\delta f(x;\Delta x) = \frac{d}{d\varepsilon} \left. f(x + \varepsilon\Delta x) \right|_{\varepsilon=0}. \tag{2.51}$$

Proof: [16]. □

These results have a straightforward generalization to Cartesian product Hilbert space.

Corollary 2.17: Suppose $f(x,y)$ is a functional over a Cartesian product Hilbert space,

$$f : \mathcal{X} \times \mathcal{Y} \to \mathbf{R}. \tag{2.52}$$

If the Fréchet differential of f exists, it is given by

$$\delta f((x,y);(\Delta x,\Delta y)) = \delta f(x,y;\Delta x) + \delta f(x,y;\Delta y) \tag{2.53}$$

where each of the differentials on the right hand side of the above equation is computed according to equation (2.51).

Proof: The proof follows from Theorem 2.16. □

Equation (2.50) can also be extended. The Riesz-Fréchet theorem implies

$$\delta f((x,y);\Delta x,\Delta y)) = \langle (\nabla_x f, \nabla_y f),(\Delta x,\Delta y) \rangle = \langle \nabla_x f,\Delta x \rangle + \langle \nabla_y f,\Delta y \rangle \tag{2.54}$$

where this last equality was from equation (2.31).

An application of Fréchet differentials that will be important in the development of this thesis is the sensitivity of eigenvalues and singular values with respect to a parameterization of a matrix. The following two theorems provide formulas for these sensitivities.

Theorem 2.18: Consider the real $n \times n$ matrix $A(\alpha)$ whose elements are Fréchet differentiable functions of the vector α,

$$A : \mathbf{R}^n \to \mathbf{R}^{n \times n}, \qquad \alpha \in \mathbf{R}^n \tag{2.55}$$

If $\lambda_i(\alpha)$ is a distinct eigenvalue of $A(\alpha)$ with right eigenvector $v_i(\alpha)$ and left eigenvector $w_i(\alpha)$ then

$$\delta\lambda_i(\alpha;\Delta\alpha) = \frac{w_i^H(\alpha)\delta A(\alpha;\Delta\alpha)v_i(\alpha)}{w_i^H(\alpha)v_i(\alpha)} . \qquad (2.56)$$

Proof: [18]. □

Although A is a matrix, it is not considered here in the usual sense of a linear operator. Each entry of A is a possibly nonlinear functional of the vector α, and hence $A(\cdot)$ is a nonlinear function from R^n into $R^{n\times n}$.

Note that the hypothesis λ_i distinct is imposed in the above theorem because for nondistinct λ_i, the left and right eigenvectors can be orthogonal, that is, $w_i^H(\alpha)v_i(\alpha) = 0$.

An analogous theorem holds for singular values.

Theorem 2.19: Consider the matrix $A(\alpha)$ as in equation (2.55). If $\sigma_i(\alpha)$ is a distinct nonzero singular value of $A(\alpha)$, then

$$\delta\sigma_i(\alpha;\Delta\alpha) = \text{Re}\{u_i^H(\alpha)\delta A(\alpha;\Delta\alpha)v_i(\alpha)\} \qquad (2.57)$$

where $u_i(\alpha)$ and $v_i(\alpha)$ are the left and right singular vectors that correspond to $\sigma_i(\alpha)$.

Proof: From equation (2.3)

$$\sigma_i^2(\alpha) = \lambda_i(A^H(\alpha)A(\alpha)). \qquad (2.58)$$

Applying the chain rule (equation (2.48)) and equation (2.46),

$$2\sigma_i(\alpha)\delta\sigma_i(\alpha;\Delta\alpha) = \frac{w_i^H(\alpha)\delta[A^H(\alpha)A(\alpha)]v_i(\alpha)}{w_i^H(\alpha)v_i(\alpha)} . \qquad (2.59)$$

Here, $w_i(\alpha)$ and $v_i(\alpha)$ are the left and right eigenvectors of the matrix

$A^H(\alpha)A(\alpha)$. Since $w_i^H(\alpha)v_i(\alpha)$ is in the denominator, we have

$$w_i^H(\alpha)v_i(\alpha) \neq 0 \rightarrow \sigma_i(\alpha) \text{ is distinct.}$$

Also, since $A^H(\alpha)A(\alpha)$ is a Hermitian matrix, we know that

$$
\begin{aligned}
&(1) \quad w_i = v_i \\
&(2) \quad v_i^H v_i = 1.
\end{aligned}
\tag{2.60}
$$

Hence,
$$w_i^H(\alpha)v_i(\alpha) = 1.$$

Applying the product rule on the right hand side of equation (2.59)

$$\delta[A^H(\alpha)A(\alpha)] = \delta A^H(\alpha;\Delta\alpha)A(\alpha) + A^H(\alpha)\delta A(\alpha;\Delta\alpha).$$

Substituting this and equation (2.60) into equation (2.59),

$$2\sigma_i(\alpha)\delta\sigma_i(\alpha;\Delta\alpha) = v_i^H(\alpha)[A^H(\alpha)\delta A(\alpha;\Delta\alpha) + \delta A^H(\alpha;\Delta\alpha)A(\alpha))v_i. \tag{2.61}$$

From equation (2.9),

$$2\sigma_i\delta\sigma_i(\alpha;\Delta\alpha) = \sigma_i(\alpha)u_i^H(\alpha)\delta A(\alpha;\Delta\alpha)v_i(\alpha) + v_i^H\delta A^H(\alpha;\Delta\alpha)u_i(\alpha)\sigma_i(\alpha)$$

where $u_i(\alpha)$ is the left singular vector of $A(\alpha)$. Then,

$$
\begin{aligned}
2\sigma_i\delta\sigma_i(\alpha;\Delta\alpha) &= \sigma_i(u_i^H\delta A(\alpha;\Delta\alpha)v_i + [u_i^H\delta A(\alpha;\Delta\alpha)v_i]^H) \\
&= 2\sigma_i \text{Re}\{u_i^H\delta A(\alpha;\Delta\alpha)v_i\}.
\end{aligned}
$$

Dividing each side by $2\sigma_i(\alpha)$ yields equation (2.57). $\qquad\qquad$ $\square$

In the σ_i nondistinct case, the singular vectors will be nonunique. However, it can be shown [19] that there exists u_i in the left singular vector subspace and v_i in the right singular vector subspace such that equation (2.57) holds.

3. ADJUSTMENT OF CONTROL SYSTEM PARAMETERS USING SINGULAR VALUE SENSITIVITIES

The purpose of this section is to develop a method for modifying control system designs to satisfy singular value inequalities. First, the translation of desirable closed loop MIMO system properties into requirements on the singular values of the return difference matrix, the loop transfer matrix, and the inverse Nyquist matrix is discussed. Then, general sensitivity formulas for these important functionals are derived. A control parameter adjustment method for manipulating the frequency shapes of these functionals is proposed. This theory is then applied to the linear quadratic (LQ) problem to obtain specific results.

3.1. Loop Shaping in the Frequency Domain for MIMO Systems

Consider the MIMO system shown in Figure 3.1, with

$$u(j\omega) \in \mathbb{C}^p$$

$$y(j\omega), \ u_c(j\omega), \ d(j\omega), \ \eta(j\omega) \in \mathbb{C}^m.$$

$P(j\omega)$ is the p input, m output plant, and $K(j\omega,\alpha)$ is the m input, p output controller. The vector α represents the control parameters, which are chosen by the designer. The vectors d and η are disturbance and sensor noise signals, respectively.

The transfer function matrix of this system is given by

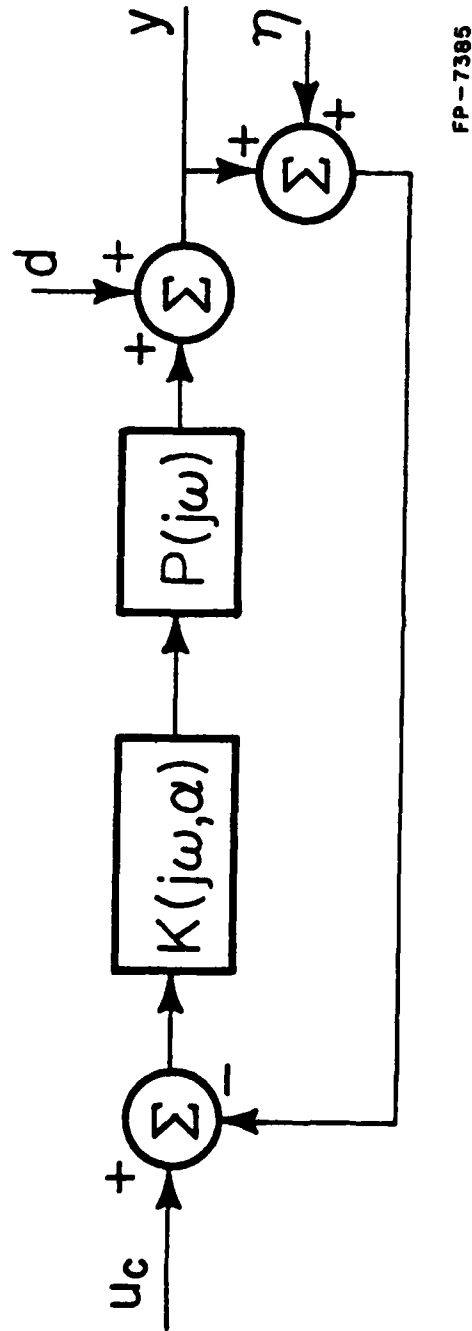$$y = PK[I + PK]^{-1}(u_c - \eta) + [I + PK]^{-1}d \qquad (3.1)$$

34

Figure 3.1. Multivariable feedback system.

where the explicit dependence on ω and α has been suppressed for notational clarity. Here, we see some of the typical SISO design tradeoffs appearing immediately in the MIMO case. For example, consider the above equation with $\eta = u_c = 0$. The effect of the disturbance on the output is

$$y = [I + PK]^{-1} d.$$

Taking norms on each side and using equations (2.10) and (2.11) we obtain the inequality

$$\| y \| \leq \frac{1}{\underline{\sigma}(I + PK)} \| d \|. \tag{3.2}$$

For disturbance rejection we must therefore have

$$\underline{\sigma}(I + PK) > 1 \tag{3.3}$$

at frequencies where disturbances may be large. This translates into a condition on the loop gains through the singular value inequality

$$\left| \underline{\sigma}(I + PK) - \underline{\sigma}(PK) \right| \leq 1. \tag{3.4}$$

When

$$\underline{\sigma}(PK) >> 1 \tag{3.5}$$

condition (3.3) is implied. This requirement of large loop gains and large return difference matrix for disturbance rejection is analagous to the scalar case.

A system has good command following properties if the error between

the input and the output is small when no sensor noise or disturbance is present. From equation (3.1), we see that good command following requires

$$PK[I + PK]^{-1} \cong I \qquad (3.6)$$

in the frequency range where the command signal has significant energy. If $e(j\omega)$ is defined

$$e(j\omega) = y(j\omega) - u_c(j\omega) \qquad (3.7)$$

then the relation

$$e = [I + PK]^{-1} u_c \qquad (3.8)$$

is obtained by substituting equation (3.1) into (3.7) with $\eta = d = 0$. As before, condition (3.3) implies that $\| e \|$ will be attenuated and that equation (3.6) will be satisfied.

The first limitation on making the loop gains arbitrarily large is imposed by the sensor noise η. From equation (3.1), the noise to output transfer function matrix is just the negative of the command to output matrix. Thus, conditions (3.5) and (3.3) and equation (3.6) imply that the sensor noise will be passed directly to the output. From equation (3.1),

$$\| y \| \leq \frac{\overline{\sigma}(PK)}{\underline{\sigma}(I + PK)} \| \eta \| \qquad (3.9)$$

and so it is desirable to have

$$\overline{\sigma}(PK) < 1 \qquad (3.10)$$

at frequencies where the command signal is dominated by the sensor noise,

There are other more severe restrictions on the loop gains imposed by robustness and sensitivity considerations.

A closed loop system is said to be robust when it remains stable under plant perturbations. These perturbations represent the difference between the true plant, $P(j\omega)$, and the mathematical model, $P^*(j\omega)$. As discussed in Section 2, this difference can be modeled by additive or multiplicative perturbations. In the case of additive perturbations that surround the plant and controller, the sufficient condition for closed loop stability (see Thm. 2.4) is

$$\underline{\sigma} \; (I + P^*(j\omega)K(j\omega,\alpha)\,) \; > \; \ell_A(j\omega) \quad \forall \; \omega \qquad (3.11)$$

where $\ell_A(j\omega)$ is the uncertainty magnitude bound.[1] For good robustness properties, the left hand side of the above equation should be made as large as possible through choice of the parameter vector α.

When the difference between $P(j\omega)$ and $P^*(j\omega)$ is modeled by a multiplicative perturbation, the sufficient condition for stability (see Thm. 2.5) is

$$\underline{\sigma}(I + (P^*(j\omega)K(j\omega,\alpha))^{-1} > \ell_M(j\omega) \quad \forall \; \omega \qquad (3.12)$$

If $\ell_M(j\omega) >> 1$, then by equation (3.4)

---

[1]The uncertainty magnitude bound $\ell_A(j\omega)$ is a positive real valued function of ω. It defines a class of perturbations ΔP_A where each member of the class is a $m \times m$ transfer function matrix that is stable, rational, and has $\|\Delta P_A\| \leq \ell_A(j\omega)$.

$$\underline{\sigma}(I + [P^*(j\omega)K(j\omega,\alpha)]^{-1}) \cong \underline{\sigma}([P^*(j\omega)K(j\omega,\alpha)]^{-1})$$

$$= [\bar{\sigma}(P^*(j\omega)K(j\omega,\alpha))]^{-1} \qquad (3.13)$$

and this reduces to condition (3.10).

Similar conditions arise from sensitivity analysis. It has been shown [5] that

$$e_c(j\omega) = [I + P(j\omega)K(j\omega,\alpha)]^{-1}e_0(j\omega) \qquad (3.14)$$

where $e_c(j\omega)$ and $e_0(j\omega)$ are defined in eqn. (2.23). This again implies that the return difference matrix (eqn. 3.3) should be large in the frequency range of interest. Notice that the true plant $P(j\omega)$ appears in the above equation. This equation can be expressed in terms of the nominal model and the perturbation magnitude bound, as shown in Theorems 2.6 and 2.7. Here, we get the conditions that

$$\underline{\sigma}(I + P^*(j\omega)K(j\omega,\alpha)) > 1 + \ell_A(j\omega) \qquad (3.15)$$

for additive perturbations, and

$$\underline{\sigma}(I + [P^*K]^{-1}) > \ell_m(j\omega) + \bar{\sigma}([P^*(j\omega)K(j\omega,\alpha)]^{-1}) \qquad (3.16)$$

for multiplicative perturbations.

In designing a control system, and in choosing α, it is necessary to know where in the frequency spectrum the uncertainty bounds are large. In general, the multiplicative uncertainty bound $\ell_M(j\omega)$ increases as frequency increases. This places a limit on the loop bandwith. The crossover frequency, defined as

$$\omega_c : \bar{\sigma}(P^*(j\omega_c)K(j\omega_c,\alpha)) = 1 \qquad (3.17)$$

must occur approximately where $\ell_M(j\omega)$ becomes greater than one. Otherwise, Theorem 2.5 will be violated. Typically additive perturbations dominate at lower frequencies, and so eqn. (3.11) should be satisfied for $\omega < \omega_c$. These requirements, in terms of loop gains, are shown in Figure 3.2. This is the MIMO generalization of classical SISO loop requirements.

The task of the designer, then, is to manipulate the singular value functionals in equations (3.11) and (3.12) through choice of the parameter vector α. In this thesis, a method of choosing α to increase (or decrease) these quantities in a gradient optimal sense is proposed.
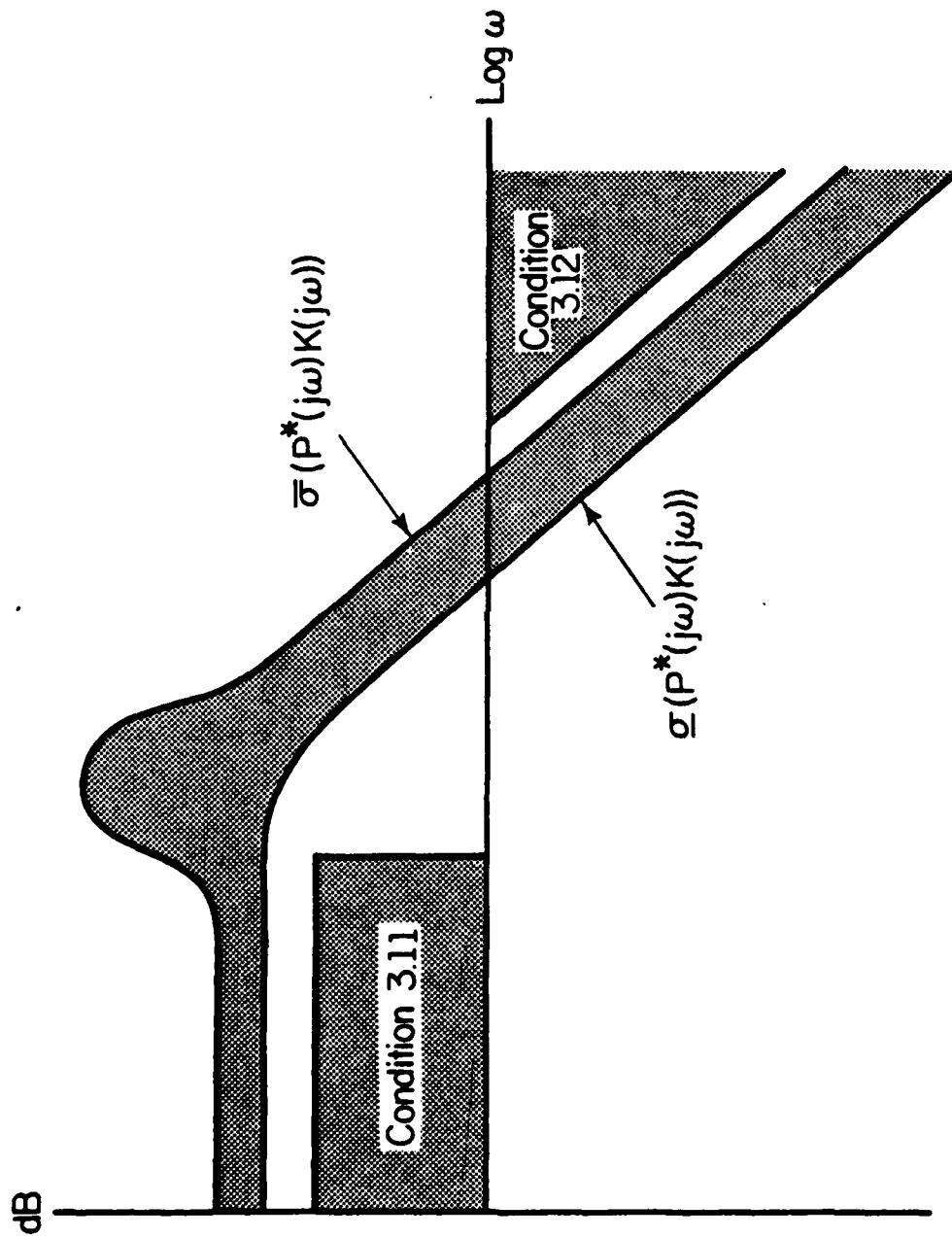
3.2. General Sensitivity Formulas

The design of any controller usually requires several iterations. An initial parameter vector α_o is selected (see Figure 3.1) and then the closed loop system is valuated to see if it meets the design objectives and constraints. If not, a new vector α_1 must be determined on the basis of the evaluation.

As previously discussed, the design objectives translate into upper and lower bounds in specified frequency ranges on the functionals defined below. In these equations, the frequency variable has been suppresed for notational brevity.

$$\underline{\sigma} \ (I + P^* K(\alpha_o)) \ \underline{\Delta} \ \underline{\sigma}_{A,0}(\alpha_o) \tag{3.18}$$

$$\underline{\sigma} \ (I + K(\alpha_o)P^*) \ \underline{\Delta} \ \underline{\sigma}_{A,I}(\alpha_o) \tag{3.19}$$

$$\underline{\sigma} \ (I + [P^* K(\alpha_o)]^{-1}) \ \underline{\Delta} \ \underline{\sigma}_{M,0}(\alpha_o) \tag{3.20}$$

Figure 3.2. Multivariable loop.

FP-7368

$$\underline{\sigma} \ (I + [K(\alpha_o)P^*]^{-1}) \ \underline{\Delta} \ \underline{\sigma}_{M,I}(\alpha_o) \qquad (3.21)$$

The next design vector α_1 must then increase (or decrease) the value of $\underline{\sigma}_{X,Y}(\alpha_o)$ in that frequency range. Here, $\underline{\sigma}_{X,Y}(\alpha_o)$ refers to any of the above functionals. (This discussion applies as well to any singular value $\sigma_{X,Y}(\alpha_o)$ of interest.)

The gradient vector $\nabla\underline{\sigma}_{X,Y}(\alpha_o)$ can be used to increase (or decrease) $\underline{\sigma}_{X,Y}(\alpha_o)$ in an optimal sense. Since the gradient vector points in the direction of maximum increase, there exists an ϵ_1 such that if $\nabla\underline{\sigma}_{X,Y}(\alpha_o) \neq 0$ then

$$\underline{\sigma}_{X,Y}(\alpha_o + \epsilon \ \frac{\nabla\underline{\sigma}_{X,Y}(\alpha_o)}{\|\nabla\underline{\sigma}_{X,Y}(\alpha_o)\|} \ ) \geq \underline{\sigma}_{X,Y}(\alpha_o + \epsilon \ \frac{u}{\|u\|}) \qquad (3.22)$$

for every vector $u \in R^n$. The vector α_1 is then defined

$$\alpha_1 \ \underline{\Delta} \ \alpha_o + \epsilon \nabla\underline{\sigma}_{X,Y}(\alpha_o). \qquad (3.23)$$

Since α_1 must be a vector of constants, a fixed value of ω must be chosen for the calculation of the gradient. The vector α_1 is then guaranteed only to increase the value of $\underline{\sigma}_{X,Y}(\alpha_o)$ in some neighborhood of $\omega = \omega_1$. To increase $\underline{\sigma}_{X,Y}(\alpha_o)$ in a specified frequency range, several design iterations may be necessary. In this case, a design iteration consists of the following procedure.

(1) Calculate $\nabla\underline{\sigma}_{X,Y}(\alpha_k)$ at a constant $\omega = \omega_k$

(2) Form $\alpha_{k+1} = \alpha_k + \epsilon\nabla\underline{\sigma}_{X,Y}(\alpha_k)$

(3) Evaluate $\underline{\sigma}_{X,Y}(\alpha_{k+1})$ as a function of frequency for different values of ϵ. Select $\epsilon = \epsilon_k$

(4) If $\underline{\sigma}_{X,Y}(\alpha_{k+1})$ does not meet the overall frequency criterion, return to step (1).

In each iteration, it may be necessary to calculate the gradient at a different frequency than the previous iteration. Also, it may be necessary to calculate gradients with respect to different functionals ($X = M,A$, $Y = I,0$).

The gradient vector can be obtained by considering

$$\underline{\sigma} : R^n \rightarrow R^+ \subset R \tag{3.24}$$

where ω is held constant. With the selection of the appropriate inner product, $\underline{\sigma}(\alpha)$ defines a map between n dimensional Hilbert space and the positive real line. When $\underline{\sigma}_{X,Y}(\alpha)$ is distinct, the Frechet differential can be computed. By the Riesz-Fréchet theorem, the differential can be expressed as an inner product. If the arbitrary vector $\Delta\alpha$ can be isolated in the inner product, the other term must be the gradient.

$$\delta\underline{\sigma}_{X,Y}(\alpha, \Delta\alpha) = \langle \nabla\underline{\sigma}_{X,Y}(\alpha), \Delta\alpha \rangle \tag{3.25}$$

This method will be used to calculate the gradient of the four singular value functionals in equations (3.18) - (3.21).

Applying the singular value differential formula (2.57) to equation (3.18),

$$\delta\underline{\sigma}_{A,0}(\alpha; \Delta\alpha) = \text{Re}\left\{ u^H \delta K(\alpha; \Delta\alpha) P^* v \right\}. \tag{3.26}$$

u and v are the left and right singular vectors of the matrix $I + P^* K(\alpha)$ that correspond to the smallest singular value. In the following equations, the symbols u and v will be used generically to denote left and right sin-

gular vectors corresponding to the appropriate matrix. Next, equation (3.19) yields

$$\delta \underline{\sigma}_{A,I}(\alpha;\Delta\alpha) = \text{Re}\left\{ u^H \delta K(\alpha;\Delta\alpha) P^* v \right\}. \qquad (3.27)$$

Similar formulas are available in the case of multiplicative perturbations. First, to avoid long equations define the loop transfer matrices

$$T_0(\alpha) \underline{\Delta} P^* K(\alpha), \qquad T_I(\alpha) \underline{\Delta} K(\alpha)P^*. \qquad (3.28)$$

Then applying equation (2.57) to (3.20),

$$\delta \underline{\sigma}_{M,0}(\alpha;\Delta\alpha) = \text{Re}\left\{ u^H \delta_\alpha \{[P^* K]^{-1}\} v \right\} \qquad (3.29)$$

where $\delta_\alpha(.)$ denotes the differential with respect to α in the direction $\Delta\alpha$. Employing equation (2.40),

$$\delta \underline{\sigma}_{M,0}(\alpha;\Delta\alpha) = -\text{Re}\left\{ u^H T_0^{-1} P^* \delta K(\alpha;\Delta\alpha) T_0^{-1} v \right\}. \qquad (3.30)$$

Similarly,

$$\delta \underline{\sigma}_{M,I}(\alpha;\Delta\alpha) = -\text{Re}\left\{ u^H T_I^{-1} \delta K(\alpha;\Delta\alpha) P^* T_I^{-1} v \right\}. \qquad (3.31)$$

To isolate the $\Delta\alpha$ term in the above equations, the structure of the controller, the choice of inner product, and the particular space to which α belongs must be determined. This will depend on the particular problem under consideration.

3.3. The Linear Quadratic Problem

The linear quadratic (LQ) control problem expresses the control objectives in terms of a performance index. The objective of the problem

is to choose a control u to minimize

$$J = \int_0^\infty [x^T Q x + u^T R u] dt \qquad (3.32)$$

where x is the state vector defined by the system

$$\dot{x} = Ax + Bu. \qquad (3.33)$$

Here,

$$x \in \mathcal{R}^n, \quad u \in \mathcal{R}^m.$$

It will be assumed that:

(1) $Q \geq 0$ (3.34)

(2) $R > 0$

(3) (A,B) is stabilizable

(4) $(\sqrt{Q}, A)$ is detectable

(5) All states are measurable

Then, a linear stabilizing state feedback solution exists and is given by

$$u = -Gx \qquad (3.35)$$

where

$$G = R^{-1} B^T K$$

and K is the unique positive semidefinite stabilizing solution of the algebraic riccati equation (ARE),

$$KA + A^T K + Q - KBR^{-1}B^T K = 0. \qquad (3.36)$$

To make this feedback system realistically feasible, a command input is usually added to the control u. Equation (3.35) then becomes

$$u = -Gx + u_c$$

The system of equation (3.33) with the control given above can be expressed in the frequency domain by transfer function matrices. The nominal plant transfer function matrix is given by

$$P^*(j\omega) = \varphi(j\omega)B \qquad (3.37)$$

where

$$\varphi(j\omega) \underline{\Delta} (j\omega I - A)^{-1}.$$

The loop transfer matrix, calculated by opening the loop at the input of the plant is

$$T_I(Q,R) = G\varphi B \qquad (3.38)$$

with

$$T_I(Q,R) \in \mathcal{C}^{m \times m}.$$

The dependence of T_I on the Q and R matrices is considered explicitly here. Also, the dependence on ω is suppressed for notational brevity. The output loop transfer matrix is

$$T_0(Q,R) = \varphi BG \qquad (3.39)$$

where

$$T_0(Q,R) \in \mathcal{C}^{n \times n}.$$

The closed loop transfer matrix is given by the following input-output

equation,

$$y = \varphi B[I + G\varphi B]^{-1} u_c. \tag{3.40}$$

The controller is simply the constant gain matrix G.

$$K(j\omega, \alpha) = G(Q,R). \tag{3.41}$$

Here, the control parameter α represents the (Q,R) matrix pair, since these matrices are the choice of the designer. The LQ system is shown in Figure 3.3.

Although G(Q,R) appears in the feedback loop instead of the feedfoward loop, the preceding concepts and theorems are still valid. It is still desirable to make

$$\underline{\sigma}(I + T_I(Q,R)) \tag{3.42}$$

as large as possible below crossover, and

$$\underline{\sigma}(I + T_I(Q,R)^{-1}) \tag{3.43}$$

as large as possible above crossover.

It turns out theoretically that the full state feedback LQ loop has very good performance and additive robustness properties. Starting with the ARE, it is possible to derive the following relation [20] provided that the open loop system has no poles that lie on the $j\omega$ axis.

$$\underline{\sigma}_R (I + T_I(Q,R)) \geq \sqrt{\lambda_{min}(R)} \ \forall \ \omega \ ; \ \forall \ \text{admissible (Q,R).} \tag{3.44}$$

The R subscript means that the singular value has been weighted by R.

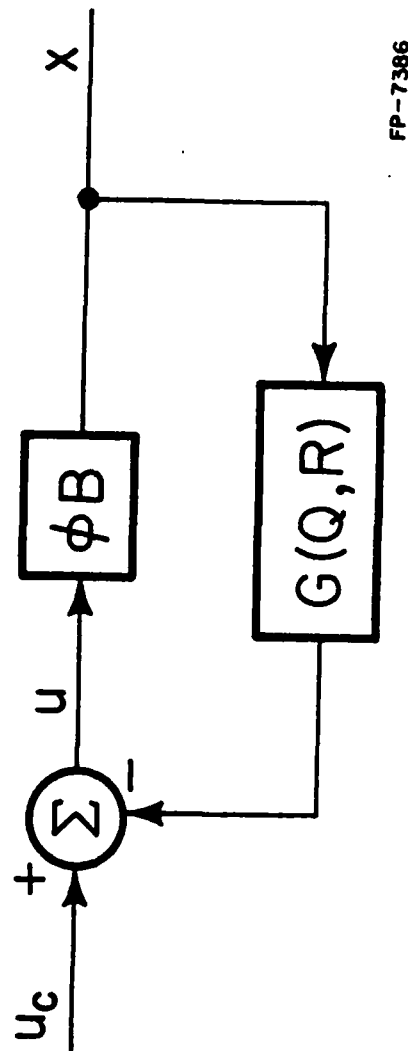$$\underline{\sigma}_R (A) \ \underline{\Delta} \ \sqrt{\lambda_{min}(A^H RA)}. \tag{3.45}$$

Figure 3.3. Linear quadratic system.

In classical terms, this translates into $+\infty$, -6dB gain margin, and $\pm 60°$ phase margin independantly in each input channel when R is a diagonal matrix [4].

Although these are impressive theoretical guarantees, in practice there are problems. The fact that every LQ loop has a 20dB/decode rolloff, contradicts the physical reality that almost every real system has a transfer function with two or more poles than zeros [2]. Inevitably, there will be unmodeled dynamics in the nominal LQ system. Hence, the designer may find it desirable to manipulate equations (3.42) and (3.43) to obtain even better properties than equation (3.44). For example, $\underline{\sigma}_{A,I} \geq 10$dB may be a necessary requirement in a certain frequency range. The iterative design method of section 3.2 provides a way to tune up the Q and R matrices to achieve the desired frequency shapes of the system singular values.

An immediate obstacle in applying this design technique to the LQ problem arises because (Q_i, R_i) must satisfy the positive definitions requirements (3.34). In equation (3.23), the parameter vector α_i represents (Q_i, R_i). For example,

$$Q_1 = Q_0 + \epsilon \nabla \underline{\sigma}_{X,Y}(Q_0, R_0).$$

The matrix Q_1 will not necessarily be positive semi definite. This problem can be circumvented by calculating the gradient with respect to the square roots $\mathcal{J}$ (Q_0, R_0), where

$$M_0^T M_0 = Q_0 \qquad M_0 \in R^{p \times n}$$

$$N_0^T N_0 = R_0 \qquad N_0 \in R^{m \times m}.$$

(3.46)

Then, the (Q_1, R_1) pair is formed

$$Q_1 = (M_o + \epsilon_1 \nabla_M \underline{\sigma}_{i,j})^T (M_o + \epsilon_1 \nabla_M \underline{\sigma}_{i,j}), \qquad (3.47)$$

$$R_1 = (N_o + \epsilon_2 \nabla_N \underline{\sigma}_{i,j})^T (N_o + \epsilon_2 \nabla_N \underline{\sigma}_{i,j}).$$

Note that only positive semidefiniteness is guaranteed, and R_1 will have to be separately tested for positive definiteness.

Before using the sensitivity formulas (3.26) - (3.31) to calculate the gradient vector, a Hilbert space and inner product must be defined. Here, the parameter vector α represents the (M,N) matrix pair. The singular value functional $\underline{\sigma}_{X,Y}(M,N)$ is defined over the cross product Hilbert space $R^{p \times n} \times R^{m \times m}$,

$$\underline{\sigma}_{X,Y} : R^{p \times n} \times R^{m \times m} \rightarrow R^+ \subset R. \qquad (3.48)$$

The inner product is defined as

$$\langle (M_1 N_1), (M_2, N_2) \rangle \underline{\Delta} \ \text{tr} \{ M_1^T M_1 \} + \text{tr} \{ N_1^T N_1 \}. \qquad (3.49)$$

This inner product is the matrix analog of the Euclidean inner product on R^n.

We are now ready to derive the gradient of the functionals $\underline{\sigma}_{X,Y}(\alpha_o)$ for the LQ controller. First, consider the case of additive perturbations:

<u>Theorem 3.1</u> (Input Perturbations)

a) The gradient of $\underline{\sigma}_{A,I}(M,N)$ with respect to M is

$$\nabla_M \underline{\sigma}_{A,I} = -MP_{1,I} \qquad (3.50)$$

where $P_{1,I}$ is the solution of the Lyapunov Equation

$$P_{1,I}\bar{A}^T + \bar{A}P_{1,I} - \text{Re}\{\Sigma_{1,I} + \Sigma_{1,I}^T\} = 0 \qquad (3.51)$$

and where:

$$\Sigma_{1,I} = \varphi B v u^H R^{-1} B^T \qquad (3.52)$$

$$\bar{A} = A - BG \qquad (3.53)$$

v, u = right and left (respectively) singular vectors of $I + T_I(M,N)$ corresponding to the smallest singular value.

b) The gradient of $\underline{\sigma}_{A,I}(M,N)$ with respect to N is

$$\nabla_N \underline{\sigma}_{A,I} = -NP_{2,I} - NGP_{1,I}G^T \qquad (3.54)$$

where:

$$P_{2,I} = \text{Re}\{\Sigma_{2,I} + \Sigma_{2,I}^T\} \qquad (3.55)$$

$$\Sigma_{2,I} = T_I v u^H R^{-1}. \qquad (3.56)$$

To simplify the notation, the explicit dependence of the Fréchet differentials on the variables $M, N, \Delta M$, and ΔN will be dropped unless needed for clarity.

Proof: From equation (3.27), we have

$$\delta_{(M,N)} \underline{\sigma}_{A,I} = \text{Re}\{u^H \delta_{(M,N)} G \varphi B v\}. \qquad (3.57)$$

The differential of G over the cross product space is equal to the sum of the differentials over the individual spaces;

$$\delta_{(M,N)} G = \delta_M G + \delta_N G. \qquad (3.58)$$

First,

$$\delta_M G = R^{-1} B^T \delta_M K \qquad (3.59)$$

where K depends on M through the Reccati equation (3.36). Employing equation (2.51) to differentiate the Reccati equation with respect to M yields

$$\delta_M K A + A^T \delta K_M + \frac{\partial}{\partial \epsilon}(M + \epsilon \Delta M)^T (M + \epsilon \Delta M)\big|_{\epsilon=0}$$

$$- \delta_M K B R^{-1} B^T K - K B R^{-1} B^T \delta_M K = 0.$$

Combining terms,

$$\delta_M K (A - BG) + (A - BG)^T \delta_M K + M^T \Delta M + \Delta M^T M = 0. \qquad (3.60)$$

This can be expressed in terms of the Lyapunov operator defined in equation (2.35) and the closed loop A matrix defined in (3.53)

$$L_{\overline{A}}(\delta_M K) = -(M^T \Delta M + \Delta M^T M). \qquad (3.61)$$

Since this operator is invertable (see equation 2.37) the above equation can be solved for $\delta_M K$

$$\delta_M K = -L_{\overline{A}}^{-1}(M^T \Delta M + \Delta M^T M). \qquad (3.62)$$

Substituting this into (3.59)

$$\delta_M G = -R^{-1} B^T L_{\overline{A}}^{-1}(M^T \Delta M + \Delta M^T M). \qquad (3.63)$$

The differential $\delta_N G$ can also be found,

$$\delta_N G = R^{-1} B^T \delta_N K + \delta_N (R^{-1}) B^T K. \qquad (3.64)$$

Using equation (2.40),

$$\delta_N(R^{-1}) = -R^{-1}\delta_N R R^{-1}$$

$$= \frac{\partial}{\partial \epsilon} R^{-1}(N + \epsilon \Delta N)^T (N + \epsilon \Delta N) R^{-1} \Big|_{\epsilon = 0}$$

$$= -R^{-1}(N^T \Delta N + \Delta N^T N) R^{-1}. \tag{3.65}$$

The differential $\delta_N K$ can be found by differentiating the Riccati equation with respect to N,

$$\delta K_N A + A^T \delta_N K - \delta_N KBR^{-1}B^T K - KB\delta_N(R^{-1})B^T K - KBR^{-1}B^T \delta_N K = 0.$$

Substituting (3.65) into the above and combining terms gives

$$\delta_N K \bar{A} + \bar{A} \delta_N K + G^T (N^T \Delta N + \Delta N^T N) G = 0.$$

In terms of $L_{\bar{A}}$ we have

$$L_{\bar{A}}(\delta_N K) = -G^T (N^T \Delta N + \Delta N^T N) G$$

$$\delta_N K = -L_{\bar{A}}^{-1} (G^T (N^T \Delta N + \Delta N^T N) G). \tag{3.66}$$

Substituting (3.66) and (3.65) into (3.64) yields

$$\delta_N G = -R^{-1}B^T L_{\bar{A}}^{-1} (G^T (N^T \Delta N + \Delta N^T N) G)$$

$$-R^{-1}(N^T \Delta N + \Delta N^T N) G. \tag{3.67}$$

Substituting equations (3.67) and (3.63) into (3.58),

$$\delta_{(M,N)}G = -R^{-1}B^TL_{\overline{A}}^{-1}(M^T\Delta M + \Delta M^TM) - R^{-1}B^TL_{\overline{A}}^{-1}(G^T(N^T\Delta N$$

$$+ \Delta N^TN)G) - R^{-1}(N^T\Delta N + \Delta N^TN)G. \tag{3.68}$$

Then, from equation (3.57),

$$\delta_{(M,N)}\underline{\sigma}_{A,I} = -\text{Re}\{u^H[R^{-1}B^TL_{\overline{A}}^{-1}(M^T\Delta M + \Delta M^TM) + R^{-1}B^TL_{\overline{A}}^{-1}(G^T(N^T\Delta N$$

$$+ \Delta N^TN)G) + R^{-1}(N^T\Delta N + \Delta N^TN)G]\varphi Bv\}. \tag{3.69}$$

Since the trace of a scalar is a scalar, we can take the trace of the
right hand side of the above equation and the equality still holds. This
allows us to manipulate the matrix terms and get the right hand side in terms
of the inner products defined on the underlying spaces. Then,

$$\delta_{(M,N)}\underline{\sigma}_{A,I} = -\text{Re}\{\text{tr}\{u^HR^{-1}(N^T\Delta N + \Delta N^TN)T_Iv\}$$

$$+ \text{tr}\{u^HR^{-1}B^TL_{\overline{A}}^{-1}(G^T(N^T\Delta N + \Delta N^TN)G)\varphi Bv\} \tag{3.70}$$

$$+ \text{tr}\{u^HR^{-1}B^TL_{\overline{A}}^{-1}(M^T\Delta M + \Delta M^TM)\varphi Bv\} \}.$$

The first term of the above can be rearranged as follows,

$$\text{First term} = \text{Re}\{\text{tr}\{u^HR^{-1}(N^T\Delta N + \Delta N^TN)T_I(M,N)v\} \}$$

$$\tag{3.71}$$

$$= \text{Re}\{\text{tr}\{u^HR^{-1}N^T\Delta NT_I(M,N)v\} + \text{tr}\{u^HR^{-1}\Delta N^TNT_I(M,N)v\}\}.$$

Using the communitive multiplication properties of the trace and definition
(3.56), we get

$$\text{First term} = \text{Re}\{\text{tr}\{\Delta N^TN(\Sigma_{2,I}^T + \Sigma_{2,I})\} \}$$

$$= \text{tr}\{\Delta N^T N \text{Re}\{\Sigma_{2,I}^T + \Sigma_{2,I}\}\}$$

$$= \text{tr}\{\Delta N^T N P_{2,I}\} \tag{3.72}$$

where $P_{2,I}$ is defined in (3.55).

The second term of (3.70) can be manipulated in a similar fashion. We have

$$\text{Second term} = \text{Re}\{\text{tr}\{u^H R^{-1} B^T L_{\bar{A}}^{-1}(G^T(N^T\Delta N + \Delta N^T N)G)\varphi Bv\}\}$$

$$= \text{Re}\{\text{tr}\{L_{\bar{A}}^{-1}(G(N^T\Delta N + \Delta N^T N)G)\varphi Bvu^H R^{-1}B^T\}\}. \tag{3.73}$$

The ΔN matrices can be exposed by moving the Lyapunov operator $L_{\bar{A}}$ on to the following terms through the adjoint operator $L_{\bar{A}}^T$ (see equation (2.38)). The above equation is then

$$\text{Second term} = \text{Re}\{\text{tr}\{G^T(N^T\Delta N + \Delta N^T N)GL_{\bar{A}}^T{}^{-1}(\varphi Bvu^H R^{-1}B^T)\}\}.$$

Breaking this up,

$$\text{Second term} = \text{Re}\{\text{tr}\{G^T N^T \Delta NGL_{\bar{A}}^T{}^{-1}(\Sigma_{1,I})\}$$

$$+ \text{tr}\{G^T\Delta N^T NGL_{\bar{A}}^T{}^{-1}(\Sigma_{1,I})\}\} \tag{3.74}$$

where $\Sigma_{1,I}$ is defined in (3.52). Using elementary trace properties,

$$\text{Re}\{\text{tr}\{GL_{\bar{A}}^T{}^{-1}(\Sigma_{1,I})G^T N^T \Delta N\} + \text{tr}\{\Delta N^T NGL_{\bar{A}}^T{}^{-1}(\Sigma_{1,I})G\}\}$$

$$= \text{Re}\{\text{tr}\{\Delta N^T NGL_{\bar{A}}^T{}^{-1}(\Sigma_{1,I} + \Sigma_{1,I}^T)G^T\}\}$$

$$= \text{tr}\{\Delta N^T NGL_{\bar{A}}^T{}^{-1}(\text{Re}\{\Sigma_{1,I} + \Sigma_{1,I}\})G^T\}\} \tag{3.75}$$

where the linear character of $L_{\underline{A}T}^{-1}$ has been exploited. According to definition (3.51), this can be written as

$$= tr\{\Delta N^T N G P_{1,I} G^T\}\}. \tag{3.76}$$

The last term of (3.70) can be manipulated similarly,

$$\text{Third term} = \{tr\{u^H R^{-1} B^T L_{\underline{A}}^{-1} (M^T \Delta M + \Delta M^T M) \varphi B v\}\}$$

$$= Re\{tr\{L_{\underline{A}}^{-1} (M^T \Delta M + \Delta M^T M) \varphi B v u^H R^{-1} B^T\}\}$$

$$= Re\{tr\{(M^T \Delta M + \Delta M^T M) L_{\underline{A}T}^{-1} (\Sigma_{1,I})\}. \tag{3.77}$$

Further juggling as in equations (3.74) - (3.75) yields

$$\text{Third term} = tr\{\Delta M^T M L_{\underline{A}T}^{-1} (Re\{\Sigma_{1,I} + \Sigma_{1,I}^T\})\}$$

$$= tr\{\Delta M^T M P_{1,I}\}. \tag{3.78}$$

Substituting equations (3.72), (3.76) and (3.78) into (3.70), we have

$$\delta_{(M,N)} \underline{\sigma}_A = -tr\{\Delta N^T N P_{2,I}\} - tr\{\Delta N^T N G P_{1,I} G^T\} - tr\{\Delta M^T M P_{1,I}\}.$$

In terms of the respective inner products,

$$\delta_{(M,N)} \underline{\sigma}_A = - \langle \Delta N, N G P_{1,I} G^T \rangle - \langle \Delta N, N P_{2,I} \rangle - \langle \Delta M, M P_{1,I} \rangle$$

$$= \langle \Delta N, -N G P_{1,I} G^T - N P_{2,I} \rangle + \langle \Delta M, -M P_{1,I} \rangle.$$

The right hand entries of these inner products are the gradient matrices. □

Theorem 3.2 (Output Perturbations)

a) The gradient of $\underline{\sigma}_{A,0}(M,N)$ with respect to M is

$$\nabla_M \underline{\sigma}_{A,0}(M,N) = -MP_{1,0} \tag{3.79}$$

where $P_{1,0}$ is the solution of the Lyapunov equation

$$P_{1,0}\overline{A}^T + \overline{A}P_{1,0} - \text{Re}\{\Sigma_{1,0} + \Sigma_{1,0}\} = 0 \tag{3.80}$$

and where: $\qquad \Sigma_{1,0} = vu^H\varphi BR^{-1}B^T \tag{3.81}$

u, v are the left and right singular vectors corresponding to the smallest singular value of the matrix $I + T_0$.

b) The gradient of $\underline{\sigma}_{A,0}(M,N)$ with respect to N is

$$\nabla_N \underline{\sigma}_{A,0}(M,N) = -NP_{2,0} - NGP_{1,0}G^T \tag{3.82}$$

where: $\qquad P_{2,0} = \text{Re}\{\Sigma_{2,0} + \Sigma_{2,0}^T\} \tag{3.83}$

$$\Sigma_{2,0} = Gvu^H\varphi BR^{-1}. \tag{3.84}$$

Proof: From equation (3.26), the differential is

$$\delta_{(M,N)}\underline{\sigma}_{A,0} = \text{Re}\{u^H\varphi B\delta_{(M,N)}Gv\}.$$

The differential of G was previously calculated in equation (3.68). Substituting this into the above, and taking the trace,

$$\delta_{(M,N)}\underline{\sigma}_{A,0} = -\text{Re}\{\text{tr}\{u^H\varphi B[R^{-1}(N^T\Delta N + \Delta N^T N)G$$

$$+ R^{-1}B^{T}L_{\underline{A}}^{-1}\{G^{T}(N^{T}\Delta N + \Delta N^{T}N)G\}$$

$$+ R^{-1}B^{T}L_{\underline{A}}^{-1}\{M^{T}\Delta M + \Delta M^{T}M]v\}\}. \tag{3.85}$$

Arranging the first term as in equation (3.71),

$$\text{First term} = \text{Re}\{\text{tr}\{u^{H}\varphi BR^{-1}(N^{T}\Delta N + \Delta N^{T}N)Gv\}\}$$

$$= \text{Re}\{\text{tr}\{u^{H}\varphi BR^{-1}N^{T}\Delta NGv\} + \text{tr}\{u^{H}\varphi BR^{-1}\Delta N^{T}NGv\}\}.$$

Using the communitive and transpose properties of the trace and definitions (3.83) and (3.84),

$$\text{First term} = \text{Re}\{\text{tr}\{Gvu^{H}\varphi BR^{-1}N^{T}\Delta N\} + \text{tr}\{\Delta N^{T}NGvu^{H}\varphi BR^{-1}\}\}$$

$$= \text{Re}\{\text{tr}\{\Sigma_{2,0}N^{T}\Delta N\} + \text{tr}\{\Delta N^{T}N\Sigma_{2,0}\}\}$$

$$= \text{tr}\{\Delta N^{T}NP_{2,0}\}. \tag{3.86}$$

The second term of (3.85) becomes

$$\text{Second term} = \text{Re}\{\text{tr}\{u^{H}\varphi BR^{-1}B^{T}L_{\underline{A}}^{-1}\{G^{T}(N^{T}\Delta N + \Delta N^{T}N)G\}v\}\}$$

$$= \text{Re}\{\text{tr}\{L_{\underline{A}}^{-1}\{G^{T}(N^{T}\Delta N + \Delta N^{T}N)G\}vu^{H}\varphi BR^{-1}B^{T}\}\}$$

$$= \text{Re}\{\text{tr}\{L_{\underline{A}}^{-1}\{G^{T}(N^{T}\Delta N + \Delta N^{T}N)G\}\Sigma_{1,0}\}\}. \tag{3.87}$$

Employing the adjoint operator and proceeding as in equation (3.74) through equation (3.77),

$$\text{Second term} = \text{tr}\{\Delta N^{T}NGP_{1,0}G^{T}\}. \tag{3.88}$$

The third term in equation (3.85) can be manipulated in a similar way,

$$\text{Third term} = \text{Re}\{\text{tr}\{u^H \varphi BR^{-1}B^T L_{\bar{A}}^{-1} \{M^T \Delta M + \Delta M^T M\}v\}$$

$$= \text{Re}\{\text{tr}\{L_{\bar{A}}^{-1}\{M^T \Delta M + \Delta M^T M\}vu^H \varphi BR^{-1}B^T\}\}$$

$$= \text{Re}\{\text{tr}\{L_{\bar{A}}^{-1}\{M^T \Delta M + \Delta M^T M\}\,\Sigma_{1,0}\}\}. \tag{3.89}$$

And proceeding as in equations (3.77) - (3.78),

$$\text{Third term} = \text{tr}\{\Delta M^T M P_{1,0}\}. \tag{3.90}$$

Substituting equations (3.90), (3.88) and (3.86) back into (3.85),

$$\delta_{(M,N)}\underline{\sigma}_{A,0} = -\text{tr}\{\Delta M^T M P_{1,0}\} - \text{tr}\{\Delta N^T N P_{2,0}\} - \text{tr}\{\Delta N^T N G P_{1,0}G^T\}$$

$$= \langle \Delta M, -MP_{1,0}\rangle + \langle \Delta N, -NP_{2,0}-NGP_{1,0}G^T\rangle.$$

This completes the proof. □

The case of multiplicative perturbations yields similar results.

Theorem 3.3 (Input Perturbation)

Suppose that the matrix B (in equation 3.33) has full column rank, and that the pair (A,M) is observable.

a) The gradient of $\underline{\sigma}_{M,I}(M,N)$ with respect to M is

$$\nabla_M \underline{\sigma}_{M,I} = MP_{3,I} \tag{3.91}$$

where $P_{3,I}$ is the solution of

$$P_{3,I}A^T \pm \bar{A}P_{3,I} - \text{Re}\{\Sigma_{3,I} + \Sigma_{3,I}^T\} = 0 \tag{3.92}$$

and where

$$\Sigma_{3,I} \triangleq \varphi B T_I^{-1} v u^H T_I^{-1} R^{-1} B^T. \tag{3.93}$$

u and v are the left and right singular vectors of $I + T_I^{-1}(M,N)$ corresponding to the smallest singular value.

b) The gradient of $\underline{\sigma}_{M,I}(M,N)$ with respect to N is

$$\nabla_N \underline{\sigma}_{M,I} = N P_{4,I} + N G P_{3,I} G^T \tag{3.94}$$

where

$$P_{4,I} \triangleq \text{Re}\{\Sigma_{4,I} + \Sigma_{4,I}^T\} \tag{3.95}$$

and,

$$\Sigma_{4,I} \triangleq v u^H T_I^{-1} R^{-1}. \tag{3.96}$$

Proof: From equation (3.31)

$$\delta_{(M,N)} \underline{\sigma}_{M,I} = -\text{Re}\{u^H T_I^{-1} \delta_{(M,N)} G \varphi B T_I^{-1} v\}. \tag{3.97}$$

The assumption that B has full rank and that (A,M) is observable is needed to guarantee the existance of the inverse of $T_I = G\varphi B$.

The differential $\delta_{(M,N)} G$ was previously computed in equation (3.68). Substituting this in,

$$\begin{aligned}
\delta_{(M,N)} \underline{\sigma}_{M,I} = \text{Re}\{u^H T_I^{-1} [R^{-1} B^T L_A^{-1} \{M^T \Delta M + \Delta M^T M\} \\
+ R^{-1} B^T L_A^{-1} \{G^T(N^T \Delta N + \Delta N^T N)G\} \\
+ R^{-1}(N^T \Delta N + \Delta N^T N)G]\varphi B T_I^{-1} v\}.
\end{aligned}$$

Taking the trace and rearranging,

$$\delta_{(M,N)}\underline{\sigma}_{M,I} = \text{Re}\{\text{tr}\{u^H T_I^{-1} R^{-1}(N^T \Delta N + \Delta N^T N)G\varphi BT_I^{-1}v\}$$

$$+ \text{tr}\{u^H T_I^{-1} R^{-1} B^T L_{\overline{A}}^{-1}\{G^T(N^T \Delta N + \Delta N^T N)G\}\varphi BT_I^{-1}v\}$$

$$+ \text{tr}\{u^H T_I^{-1} R^{-1} B^T L_{\overline{A}}^{-1}\{M^T \Delta M + \Delta M^T M\}\varphi BT_I^{-1}v\}\}. \tag{3.98}$$

The first term of the above can be manipulated as in equations (3.71)-(3.72) to yield

$$\text{Re}\{\text{tr}\{u^H T_I^{-1} R^{-1}(N^T \Delta N + \Delta N^T N)v\}\} = \text{tr}\{\Delta N^T P_{4,I}\}. \tag{3.99}$$

Following equations (3.74)-(3.76), the second term of (3.98) becomes

$$\text{Re}\{\text{tr}\{u^H T_I^{-1} R^{-1} B^T L_{\overline{A}}^{-1}\{G(N^T \Delta N + \Delta N^T N)G\}\varphi BT_I^{-1}v\}\}$$

$$= \text{tr}\{\Delta N^T NGP_{3,I}G^T\}. \tag{3.100}$$

Following equations (3.77)-(3.78) the last term of (3.98) is written

$$\text{Re}\{\text{tr}\{u^H T_I^{-1} R^{-1} B^T L_{\overline{A}}^{-1}\{M^T \Delta M + \Delta M^T M\}H\varphi BT_I^{-1}v\}\}$$

$$= \text{tr}\{\Delta M^T MP_{3,I}\}. \tag{3.101}$$

Substituting equations (3.99)-(3.101) into equation (3.98) yields

$$\delta_{(M,N)}\underline{\sigma}_{M,I} = \text{tr}\{\Delta M^T MP_{3,I}\} + \text{tr}\{\Delta N^T(NP_{4,I} + NGP_{3,I}G)\}$$

$$= \langle \Delta M, MP_{3,I}\rangle + \langle \Delta N, NP_{4,I} + NGP_{3,I}G^T\rangle.$$

This completes the proof. □

Theorem 3.4 (Output Perturbations)

Suppose that the pair (A,M) is observable, and that the matrix B has

rank n.[2]

a) The gradient of $\underline{\sigma}_{M,0}$ (M,N) with respect to M is

$$\nabla_M \underline{\sigma}_{M,0} = MP_{3,0} \tag{3.102}$$

where $P_{3,0}$ is the solution of the Lyapunov equation

$$P_{3,0}\bar{A}^T + \bar{A}P_{3,0} - Re\{\Sigma_{3,0} + \Sigma_{3,0}^T\} = 0 \tag{3.103}$$

and where:

$$\Sigma_{3,0} \triangleq T_0^{-1}vu^H T_0^{-1}\varphi BR^{-1}B^T \tag{3.104}$$

u, v are the left and right singular vectors of the matrix $I + T_0^{-1}$(M,N) corresponding to the smallest singular value.

b) The gradient of $\underline{\sigma}_{M,0}$(M,N) with respect to N is

$$\nabla_N \underline{\sigma}_{M,0} = NP_{4,0} + NGP_{3,0}G^T \tag{3.105}$$

where

$$P_{4,0} \triangleq Re\{\Sigma_{4,0} + \Sigma_{4,0}^T\} \tag{3.106}$$

and

$$\Sigma_{4,0} \triangleq GT_0^{-1}vu^H T_0^{-1}\varphi BR^{-1}. \tag{3.107}$$

Proof: From equation (3.30),

$$\delta_{(M,N)} \underline{\sigma}_{M,0} = -Re\{u^H T_0^{-1}\varphi B\delta_{(M,N)}GT_0^{-1}v\} \tag{3.108}$$

Notice that B and G are of rank n, thus the inverse of $T_0 = \varphi BG$ exists. Substituting equation (3.68) into the above and taking the trace,

---

[2]This assumption is unrealistic. See Chapter 5 for further remarks.

$$\delta_{(M,N)} \underline{\sigma}_{M,0} = \text{Re}\{\text{tr}\{u^H T_0^{-1} \varphi B [R^{-1}(N^T \Delta N + \Delta N^T N)G$$

$$+ R^{-1} B^T L_A^{-1} \{G^T (N^T \Delta N + \Delta N^T N)G\}$$

$$+ R^{-1} B^T L_A^{-1} \{M^T \Delta M + \Delta M^T M\}] T_0^{-1} v\}\}. \tag{3.109}$$

The first term of the above can be rearranged as in equations (3.71)-(3.72)

$$\text{Re}\{\text{tr}\{u^H T_0^{-1} \varphi B R^{-1}(N^T \Delta N + \Delta N^T N)G T_0^{-1} v\}\} = \text{tr}\{\Delta N^T N P_{4,0}\}. \tag{3.110}$$

The second term of equation (3.109) can be rearranged as in equations (3.74)-(3.76)

$$\text{Re}\{\text{tr}\{u^H T_0^{-1} \varphi B R^{-1} B^T L_A^{-1} \{G^T (N^T \Delta N + \Delta N^T N)G\} H T_0^{-1} v\}\}$$

$$= \text{tr}\{\Delta N^T N G P_{3,0} G^T\}. \tag{3.111}$$

Following equations (3.77)-(3.78) the last term of (3.109) becomes

$$\text{Re}\{\text{tr}\{u^H T_0^{-1} \varphi B R^{-1} B^T L_A^{-1} \{M^T \Delta M + \Delta M^T M\} H T_0^{-1} v\} = \text{tr}\{\Delta M^T M P_{3,0}\}. \tag{3.112}$$

Substituting equations (3.110)-(3.112) into equation (3.109) gives us the final form,

$$\delta_{(M,N)} \underline{\sigma}_{M,0} = \langle \Delta M, M P_{3,0} \rangle + \langle \Delta N, N P_{4,0} + N G P_{3,0} G^T \rangle$$

which proves the theorem. □

In many LQ designs an observer must be used to get full state feedback. It has been shown that the inclusion of an observer (or any dynamic element) in the LQ loop negates the robustness guarantees implied by equation (3.44),[21]. Hence the need to manipulate the system singular values

to obtain increased robustness and sensitivity margins is critical in the output feedback case.

The calculation of the gradient vector is complicated slightly by the inclusion of an observer in the loop. The observer transfer function matrix also depends on the feedback gain matrix G, and hence on the (M,N) matrix pair. The product rule for differentiation must then be used in equations (3.26) - (3.31). Following this line of reasoning, it is a straightfoward exercise to extend the above theorems to the output feedback problem.

4. APPLICATION

In this section the gradient design method will be applied to a 4th order nominal linear quadratic regulator to improve the low frequency properties. An alternative asymptotic design method which has been successfully used [22] will also be applied. The two designs will be compared, and certain important advantages of the gradient design will be discussed. The nominal plant consists of the standard linear time invariant differential equations in state variable form

$$\dot{x} = Ax + Bu, \qquad x \in R^n, \quad u \in R^m. \tag{4.1}$$

The feedback control which minimizes the cost functional

$$J = \int_0^\infty x^T Q x + u^T R u \, dt \tag{4.2}$$

is given by

$$u = -Gx \tag{4.3}$$

where

$$G = R^{-1} B^T K.$$

K is the solution of the A.R.E.,

$$KA + A^T K + Q - KBR^{-1}B^T K = 0 \tag{4.4}$$

where it is assumed that (A,B) is stabilizable and $(\sqrt{Q},A)$ is detectable. Q and R satisfy the usual positive definiteness requirements, and otherwise are arbitrary.

To illustrate the design procedure, values for A, B, Q, and R were selected as shown below

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -50 & -87 & -45 & -9 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (4.5)$$

$$Q_o = I_4 \qquad\qquad R_o = I_2. \qquad (4.6)$$

The subscript $_o$ above indicates initial values for the weighting matrices.
The nominal plant is stable with eigenvalues at

$$\lambda_1 = -1$$
$$\lambda_2 = -2 \qquad\qquad (4.7)$$
$$\lambda_{3,4} = -3 \pm j4.$$

The closed loop system has eigenvalues at

$$\lambda_{1,2} = -8.01 \pm j8.24$$
$$\lambda_{3,4} = -1.16 \pm j.408 \qquad (4.8)$$

with

$$G_o = \begin{bmatrix} 2.89 & 4.13 & 2.13 & -0.161 \\ 2.38 & 4.45 & 1.59 & -0.128 \end{bmatrix}. \qquad (4.9)$$

As discussed in Section 3.1, the singular values of the return
difference matrix, the loop matrix, and the inverse Nyquist matrix provide
a measure of the various closed loop system properties (robustness, sensi-
tivity, disturbance rejection, etc.). These quantities are plotted in
Figures 4.1, 4.2, and 4.3 respectively, for this nominal system. For con-
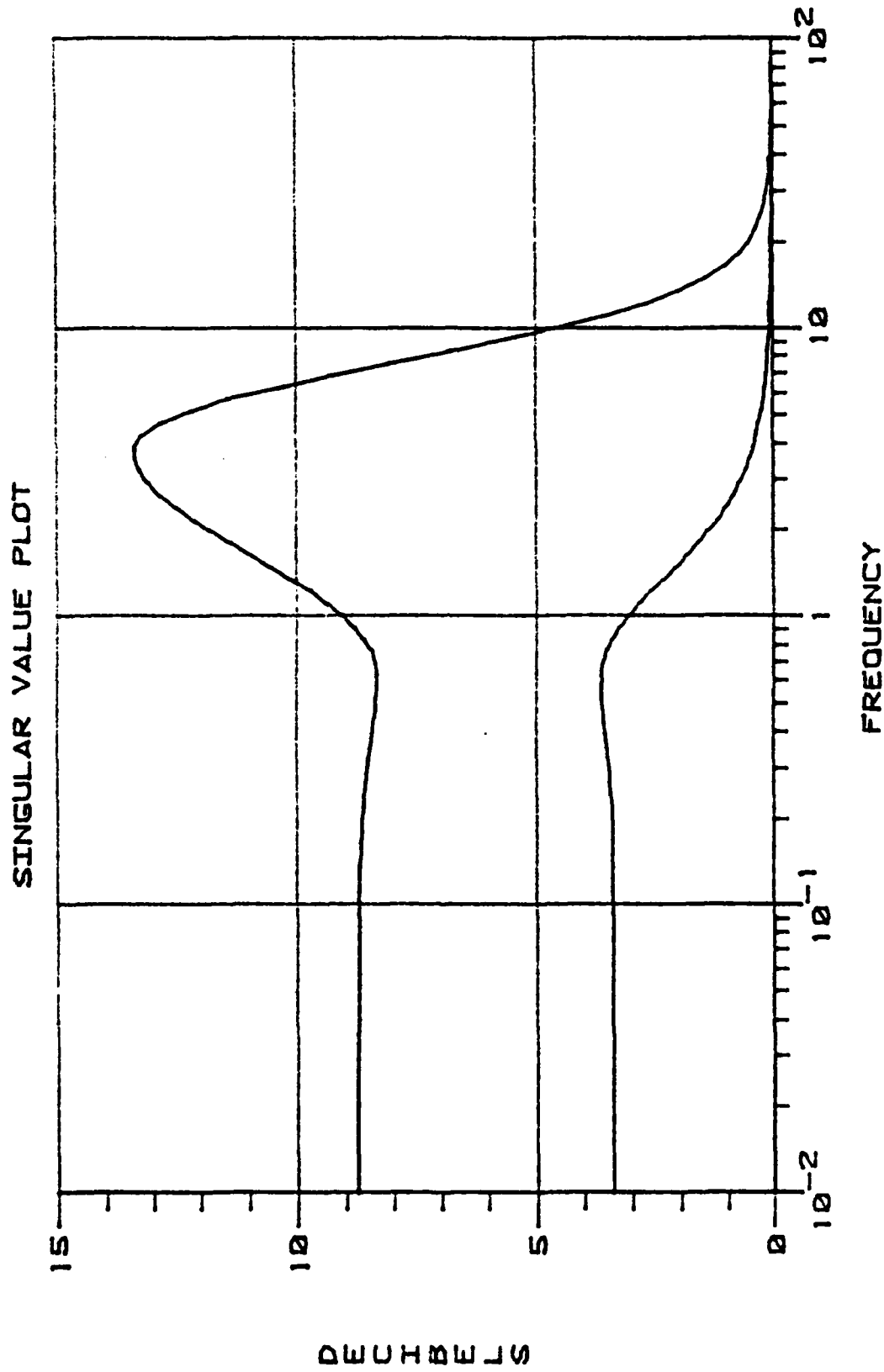venience, the notation

Figure 4.1. Singular values of the return difference matrix for the nominal system.
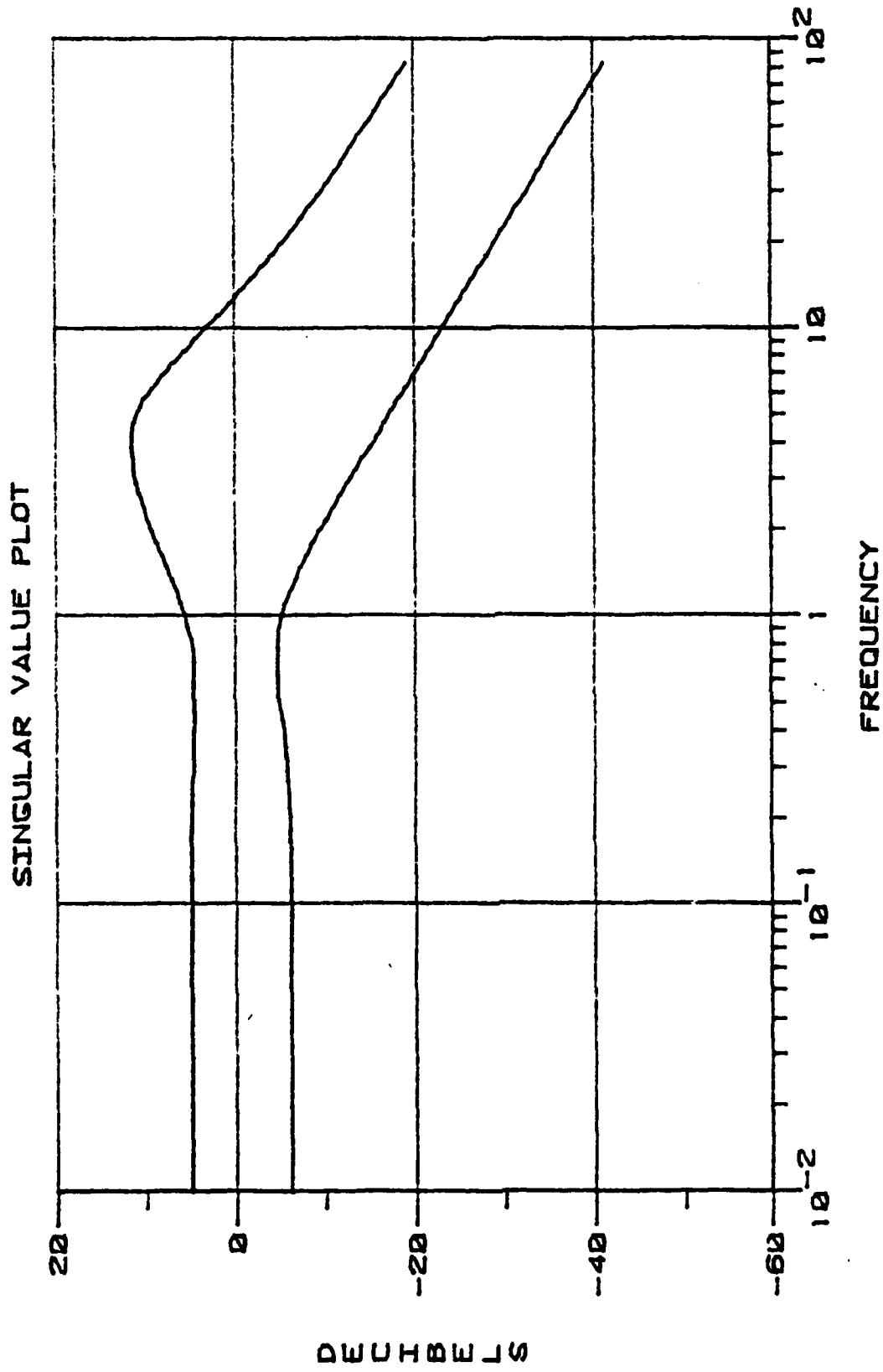
SINGULAR VALUE PLOT

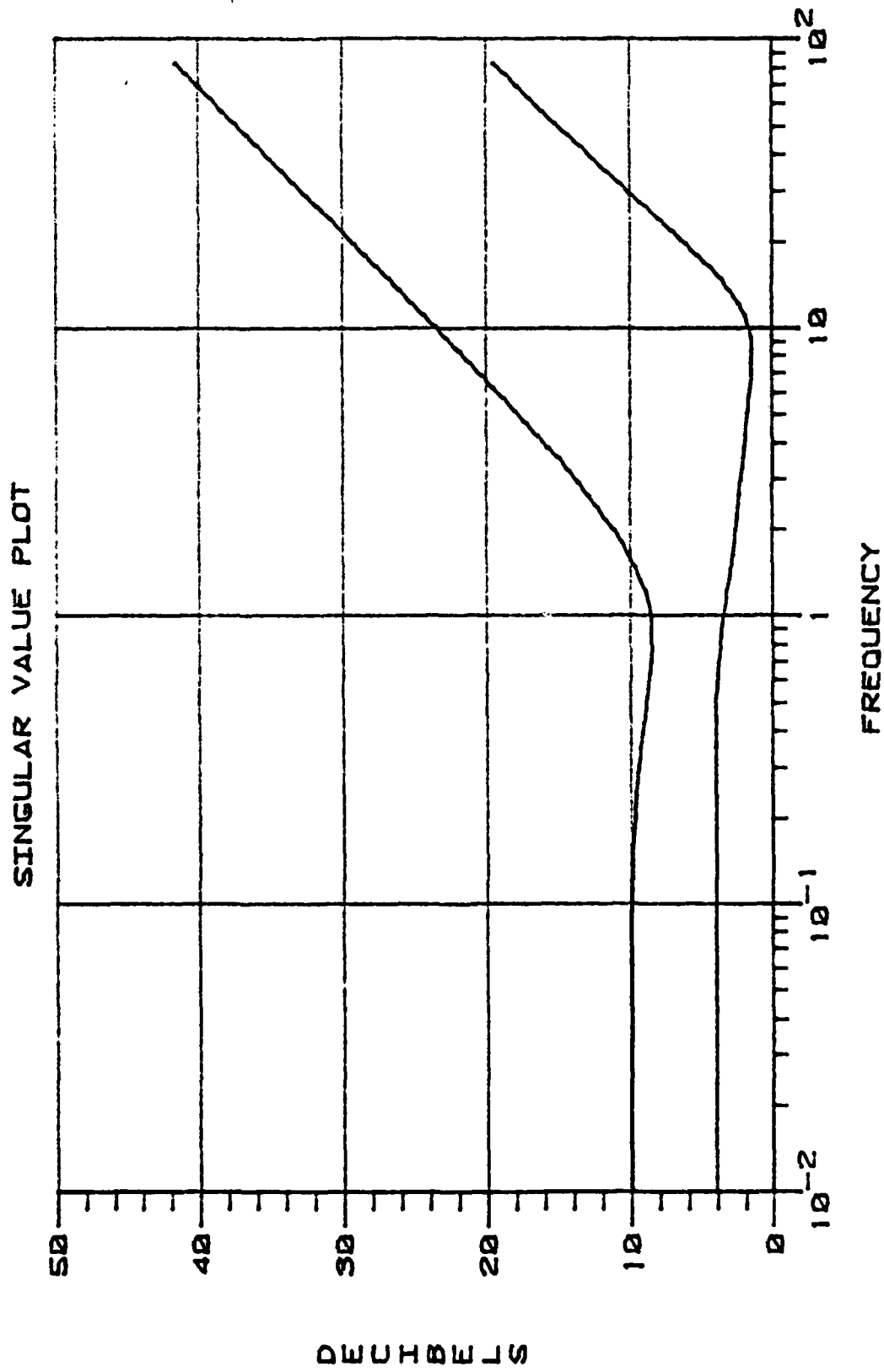Figure 4.2. Singular values of the loop transfer matrix for the nominal system.

Figure 4.3. Singular values of the inverse Nyquist matrix for the nominal system.

$$\sigma(I + G_i \varphi B) \overset{\Delta}{=} \sigma_{A,I,i} \tag{4.10}$$

$$\sigma(G_i \varphi B) \overset{\Delta}{=} \sigma_{L,I,i} \tag{4.11}$$

$$\sigma(I + [G_i \varphi B]^{-1}) \overset{\Delta}{=} \sigma_{M,I,i} \tag{4.12}$$

will be used. The i subscript refers to the independent variables (Q_i, R_i).

In this thesis, the bandwidth (ω_B) of a MIMO system will be defined to be the frequency at which the largest singular value of the transfer function matrix drops to $1/\sqrt{2}$ of its zero frequency value. Recalling that

$$\bar{\sigma}([I + G_i \varphi B]^{-1} G_i B) = [\underline{\sigma}(I + [G_i \varphi B]^{-1})]^{-1} \tag{4.13}$$

the closed loop bandwidth can l found from the plot of $\sigma_{M,I,i}$. From Figure 4.3, the closed loop bandwidth is $\omega_B \simeq 21$ rad/sec. Notice that here we are considering the input-output response between u_c and y, as shown in Figure 4.4.

The crossover frequency (ω_c) is defined to be the frequency at which the largest singular value of the loop transfer matrix has magnitude one. From Figure 4.2 $w_c \cong 11$ rad/sec.

In most LQ designs, the (Q,R) matrices are fine tuned by an iterative trial and error method. This is because there are usually several different design objectives and many constraints. Since there is no direct relationship between (Q,R) and these objectives and constraints, a series of (Q_i, R_i) matrices must be chosen, and the closed loop system tested at each iteration. The information obtained at the ith analysis is used to make the next adjustment on (Q_i, R_i). In the simple example presented here, only one iteration is necessary to expose the characteristics of the design method developed in Section 3.

For purposes of illustration, the design objective in this example will be simply to increase the value of $\underline{\sigma}_{A,I,0}$ at frequencies $\omega < .1$. This will reduce the effect of random disturbances in this frequency range. It will also make the system more robust and less sensitive to additive modeling errors (in the same frequency range) as shown in Figure 4.4. Notice that the perturbation surrounds the plant and controller. In terms of Theorem 2.4, the nominal plant includes the controller,

$$P^*(j\omega) = G_1 \varphi B. \tag{4.14}$$

In this case, the value of $\underline{\sigma}_{A,I}$ is a nonconservative bound for robustness. In other words, there exists a perturbation as shown in Figure 4.4 with

$$\|\Delta P_A\| = \underline{\sigma}_{A,I,0}$$

that destabilizes the system.

For design constraints, we will assume that the nominal bandwidth and input-output properties are acceptable, and should not be altered drastically. In particular, the closed loop bandwidth should not increase much, because if it does the system becomes more susceptible to high frequency multiplicative perturbations.

One way to affect the singular values is to assume the (Q_o, R_o) matrix pair reflects the appropriate priorities. Then, the feedback is determined by trading off between control energy and speed of response. This is accomplished by introducing a positive scaling parameter in the cost,

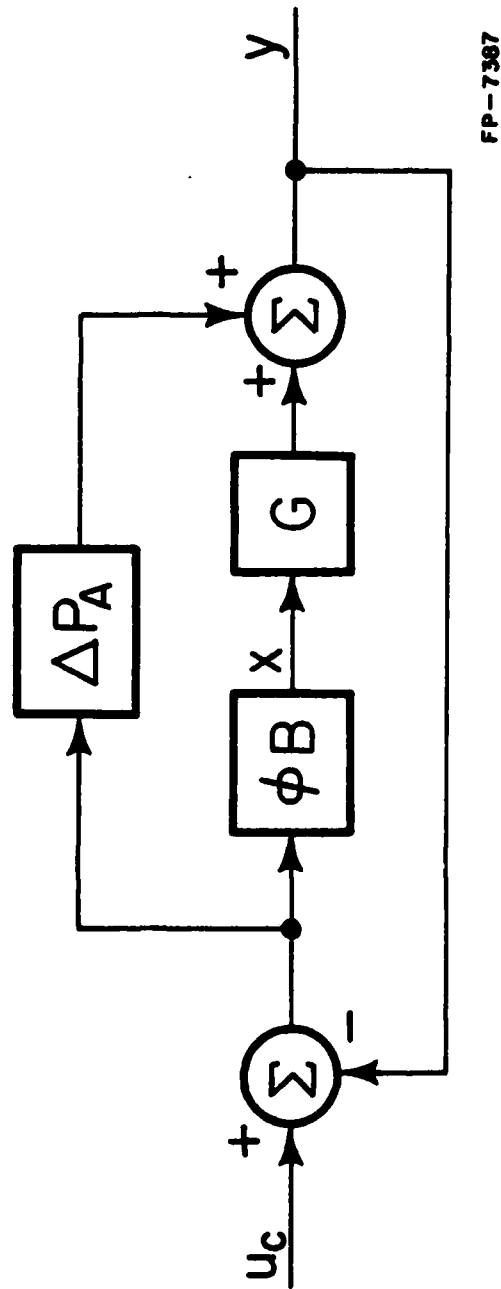$$J(\rho) = \int_0^\infty x^T Q_o x + \rho u^T R_o u \, dt. \tag{4.15}$$

Figure 4.4. Configuration of example system.

FP-7387

This method has been used for eigenvalue adjustments [10]. For large values of ρ, $J(\rho)$ approximates the minimim energy cost functional. In this case the stable open loop eigenvalues remain almost unchanged in the closed loop system. The unstable ones are reflected symmetrically about the $j\omega$ axis. As ρ tends to zero, the closed loop eigenvalues either tend to transmission zeros, or to minus infinity in a Butterworth pattern. This result is a direct consequence of the following relation which is true for any LQ system [11]

$$(I + G\varphi(j\omega)B)^H R_o(I + G\varphi(j\omega)B) = R_o + \frac{1}{\rho} [\varphi(j\omega)B]^H Q_o \varphi(j\omega)B. \qquad (4.16)$$

This equation can also be used to derive relationships between ρ and the singular value functionals.

Lemma 4.1: Consider the LQ system described in equations (4.1)-(4.4), where

$$\begin{aligned} Q &= Q_o \\ R &= \rho R_o. \end{aligned} \qquad (4.17)$$

Then, for each frequency ω for which $j\omega$ is not an eigenvalue of the open loop system,

$$\sqrt{\lambda_{min}(R_o)} \leq \lim_{\rho \to \infty} \sigma_{i_{R_o}} (I + G(\rho)\varphi B) \leq \sqrt{\lambda_{max}(R_o)} \qquad (4.18)$$

$$i = 1,2,\ldots,m$$

where $\sigma_{i_{R_o}}(\cdot)$ is the ith R_o weighted singular value, as defined in equation equation (3.45).

Proof: Consider equation (4.16). Since these matrices are equal, their eigenvalues must be equal,

$$\lambda_i((I + G(\rho)\varphi B)^H R_o(I + G(\rho)\varphi B)) = v_i^H(R_o + \frac{1}{\rho} [\varphi B]^H Q_o \varphi B)v_i. \qquad (4.19)$$

Here, v_i is the orthonormal eigenvector defined by

$$(R_o + \frac{1}{\rho} [\varphi B]^H Q_o \varphi B) v_i = \lambda_i v_i. \tag{4.20}$$

In terms of singular values and Euclidean norms

$$\sigma_{i_{R_o}} (I + G(\rho)\varphi B)^2 = \| \sqrt{R_o} v_i \|^2 + \frac{1}{\rho} \| \sqrt{Q_o} \varphi B v_i \|^2. \tag{4.21}$$

Notice that the term $\| \sqrt{Q_o} \varphi B v_i \|$ is bounded for all ρ whenever $j\omega$ is not an eigenvalue of the open loop system. Taking the limit on each side of the above equation and recalling that

$$\| v_i \| = 1 \; \Rightarrow \; \underline{\sigma}(\sqrt{R_o}) \le \| \sqrt{R_o} v_i \| \le \bar{\sigma}(\sqrt{R_o})$$

implies equation (4.18). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Notice that when $R_o = I$, equation (4.18) reduces to

$$\lim_{\rho \to \infty} \sigma_i (I + G(\rho)\varphi B) = 1.$$

A similar type of result is available when ρ tends to zero.

Lemma 4.2: Consider the LQ system described in the above lemma. Then for all ω such that $j\omega$ is not a transmission zero of $\sqrt{Q}\varphi(j\omega)B$,

$$\lim_{\rho \to 0} \underline{\sigma}(I + G(\rho)\varphi B) = \infty. \tag{4.22}$$

Proof: From equation (4.21),

$$\underline{\sigma}_{R_o} (I + G(\rho)\varphi B) \ge \frac{1}{\rho} \| \sqrt{Q_o} \varphi B v_1 \|.$$

Since $j\omega$ is not a transmission zero of $\sqrt{Q_o} \varphi B$,

$$\sqrt{Q_o} \varphi B v_1 \ne 0.$$

Consequently,

$$\lim_{\rho \to 0} \underline{\sigma}_{R_o} (I + G(\rho)\varphi B) = \infty.$$

But,

$$\underline{\sigma}_{R_o} (I + G(\rho)\varphi B) \leq \bar{\sigma}(\sqrt{R_o})\underline{\sigma}(I + G(\rho)\varphi B)$$

which implies equation (4.22).   □

If $j\omega$ is a transmission zero of $\sqrt{Q}\varphi(j\omega)B$ of multiplicity ℓ then it is still possible to see from equation (4.21) that

$$\lim_{\rho = 0} \sigma_{\ell+1}(I + G(\rho)\varphi B) = \infty. \tag{4.23}$$

The remaining ℓ singular values will be constant,

$$\sigma_{i_{R_o}} (I + G(\rho)\varphi B) = \bar{\sigma}_{R_o}(v_i) \qquad i = 1,\ldots,\ell.$$

Note that if either B or Q has rank less than m (where m is the number of inputs) then

$$\sigma_{i_{R_o}} (I + G(\rho)\varphi B) = \bar{\sigma}_{R_o}(v_i) \qquad \forall \rho$$

$$i = 1,\ldots,m - \min\{rank\{Q\}, rank\{B\}\}.$$

It is also possible to show that the R_o scaled singular values are nondecreasing functions as ρ tends to zero.

Lemma 4.3: Consider the LQ system described in equations (4.1)-(4.4) with (Q,R) as in (4.17). Then

$$\rho_1 \geq \rho_2 \quad \Rightarrow \quad \sigma_{i_{R_o}} (I + G(\rho_1)\varphi B) \leq \sigma_{i_{R_o}} (I + G(\rho_2)\varphi B) \tag{4.24}$$

where $i = 1,\ldots,m$.

__Proof:__ From equation (4.19),

$$\sigma_{i_{R_o}} (I + G(\rho)\varphi B)^2 = \lambda_1 (R_o + \frac{1}{\rho} [\varphi B]^H Q_o \varphi B).$$

Differentiating each side and utilizing equation (2.56),

$$2\sigma_{i_{R_o}} \frac{d\sigma_{i_{R_o}}}{d\rho} = - \frac{1}{\rho^2} v_1^H [\varphi B]^H Q_o \varphi B v_1.$$

Where the independent variable on the left side has been dropped for brevity.
The vector v_i is defined in equation (4.20). Then,

$$\frac{d\sigma_{i_{R_o}}}{d\rho} = - \frac{1}{2\sigma_{i_{R_o}} \rho^2} \| \sqrt{Q_o} \varphi B v_1 \|^2$$

which implies

$$\frac{d\sigma_{i_{R_o}}}{d\rho} \leq 0 \qquad \forall \rho > 0.$$

This implies equation (4.24).   □

These facts show that by scaling R_o by a parameter ρ as in
equation (4.15) the value of $\underline{\sigma}_{A,I}$ in the LQ system can be adjusted to any
desired value, as long as $\sqrt{Q_o}\varphi B$ has full rank and has no transmission zeros
on the $j\omega$ axis.

This asymptotic technique was applied to the nominal system
described in equations (4.5)-(4.6). Choosing $\rho = .26$,[1]

$$(Q_1, R_1) = (I_4, .26 I_2). \tag{4.25}$$

---

[1]This value was chosen for purposes of comparison with the following
gradient design.

This places the closed loop eigenvalues at

$$\lambda_1 = -.992$$

$$\lambda_2 = -2.32$$

$$\lambda_{3,4} = -11.3 \pm 11.3$$

with a feedback gain matrix

$$G_1 = \begin{bmatrix} 5.40 & 7.18 & 3.83 & -.563 \\ 4.28 & 8.13 & 2.33 & -.428 \end{bmatrix}. \qquad (4.26)$$

Graphs of the important quantities $\sigma_{A,I,1}$, $\sigma_{L,I,1}$, and $\sigma_{M,I,1}$ are shown in Figures 4.5, 4.6, and 4.7. From Figure 4.5, the low frequency value of $\underline{\sigma}_{A,I,1}$ is greater than $\underline{\sigma}_{A,I,0}$ by about 4dB.

This indicates an improvement in robustness, disturbance rejection and sensitivity reduction properties. However, the design constraints have been compromised somewhat. From Figure 4.6 we see that the crossover frequency has increased from the nominal value of $\omega_c \doteq 11$ rad/sec to $\omega_c \doteq 20$ rad/sec. From Figure 4.7, the bandwidth has increased from $\omega_B = 21$ rad/sec to $\omega_B = 30$ rad/sec. This shows that the system has become more susceptible to high frequency multiplicative perturbations. Also, the magnitude of the feedback matrix has increased. In the nominal system,

$$\| G_o \| = 7.12. \qquad (4.27)$$

For the asymptotically designed system,

$$\| G_1 \| = 12.8. \qquad (4.28)$$

This is an 80% increase in magnitude. It indicates that more control energy is needed to attain the increased value of $\underline{\sigma}_{A,I,1}$.
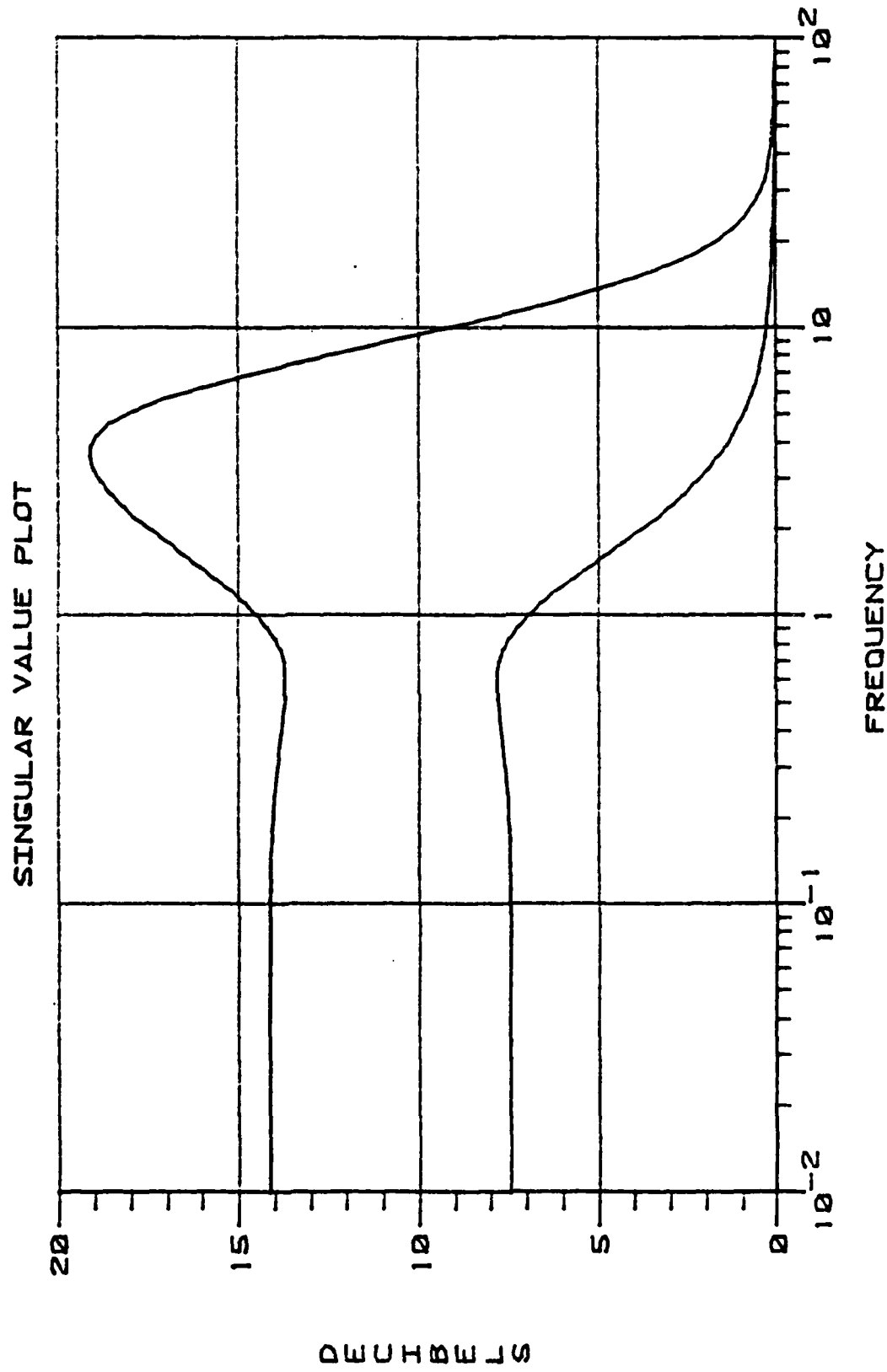
Figure 4.5. Singular values of the return difference matrix for the asymptotically designed system.
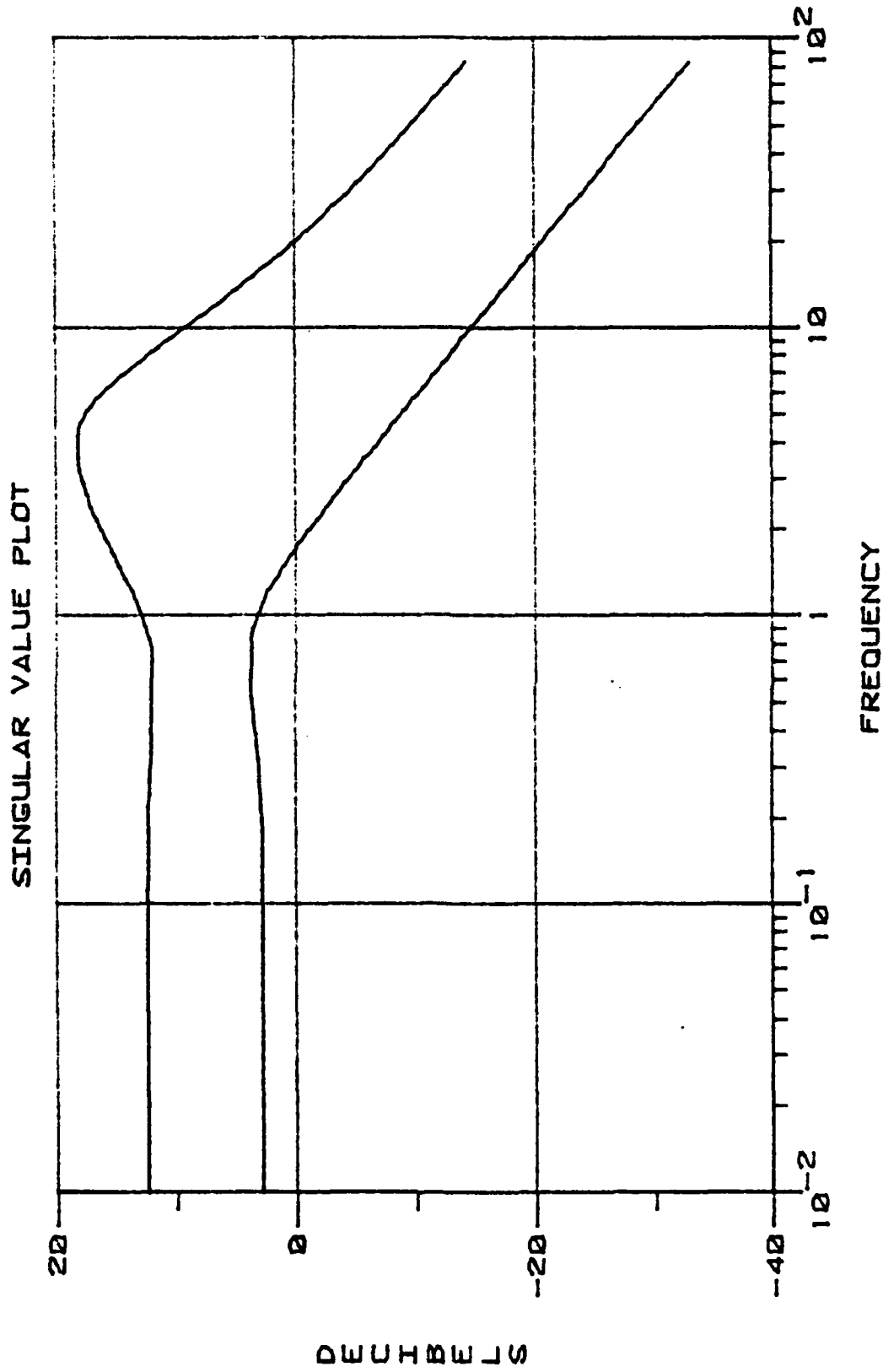
Figure 4.6. Singular values of the loop transfer matrix for the asymptotically designed system.
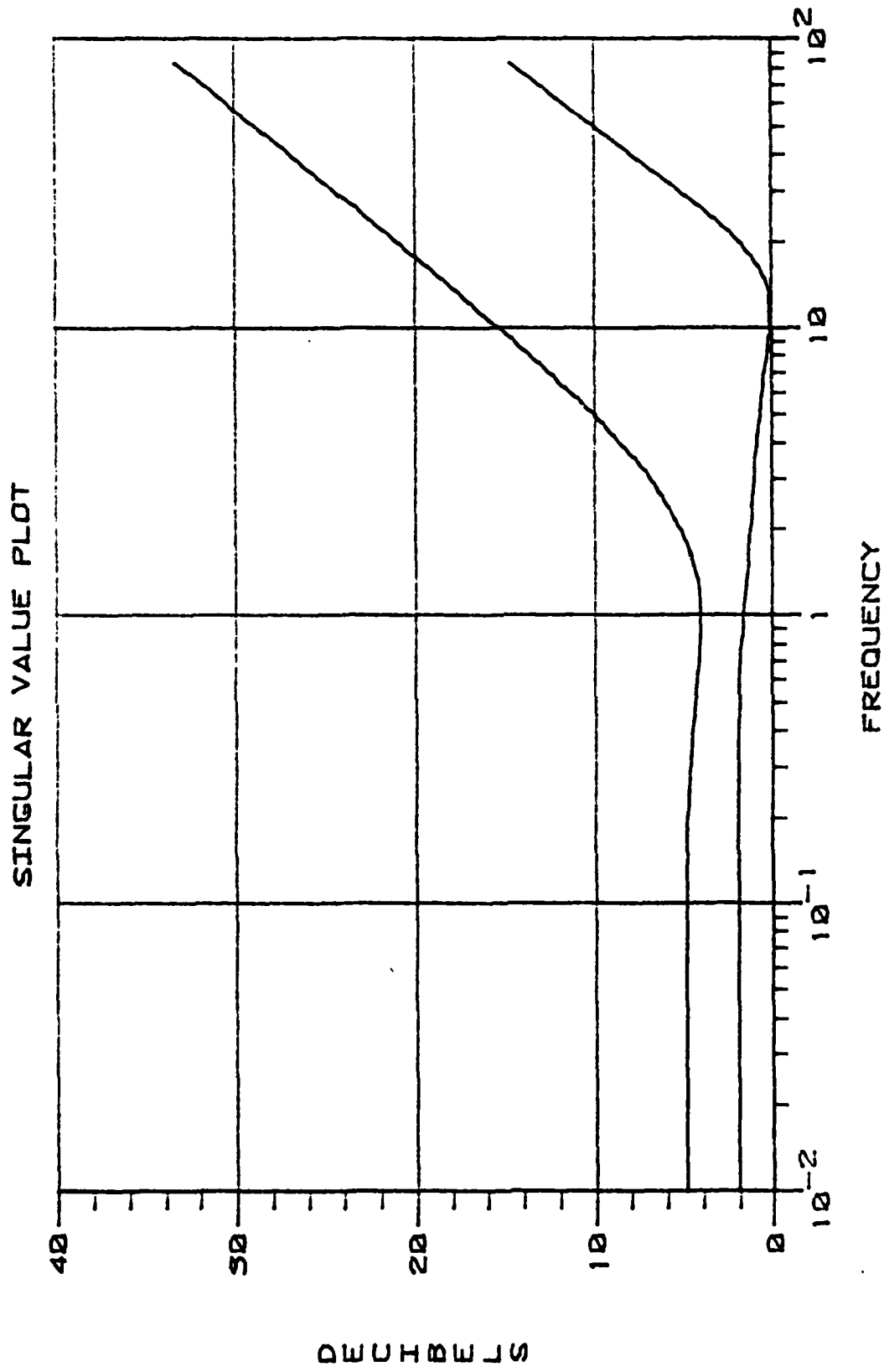
Figure 4.7. Singular values of the inverse Nyquist matrix for the asymptotically designed system.

The design procedure developed in Section 3 was also applied to this example via Theorem 3.1. The gradient of $\sigma_{A,I,0}$ was computed with respect to $M = \sqrt{Q}$, where R was considered a constant parameter. At $\omega = .1$, the gradient matrix was found to be

$$\nabla_M \sigma_{A,I,0} = \begin{bmatrix} 0.200 & 0.512 & 0.313 & -0.0525 \\ 0.512 & 1.23 & 0.131 & -0.0796 \\ 0.313 & 0.131 & 1.19 & -0.0640 \\ -0.0525 & -0.0796 & -0.0640 & -0.0281 \end{bmatrix}. \tag{4.29}$$

The Q_2 matrix was calculated from equation (3.47) with $\epsilon = 1$,

$$Q_2 = \begin{bmatrix} 1.80 & 1.80 & 1.13 & -.178 \\ 1.80 & 5.26 & .743 & -.295 \\ 1.13 & .743 & 4.91 & -.233 \\ -.178 & -.295 & -.233 & 1.07 \end{bmatrix} \tag{4.30}$$

$$R_2 = I_2.$$

It was found by trial and error that this value of ϵ produced the greatest increase of $\sigma_{A,I,2}$. Plots of $\sigma_{A,I,2}$, $\sigma_{L,I,2}$, and $\sigma_{M,I,2}$ are shown in Figures 4.8, 4.9, and 4.10. The closed loop eigenvalues are now at

$$\lambda_1 = -.871$$
$$\lambda_2 = -2.46$$
$$\lambda_{3,4} = -8.48 \pm j8.51$$

with

$$G_2 = \begin{bmatrix} 3.03 & 4.31 & 2.64 & -0.167 \\ 2.84 & 5.80 & 1.31 & -0.181 \end{bmatrix} \tag{4.31}$$
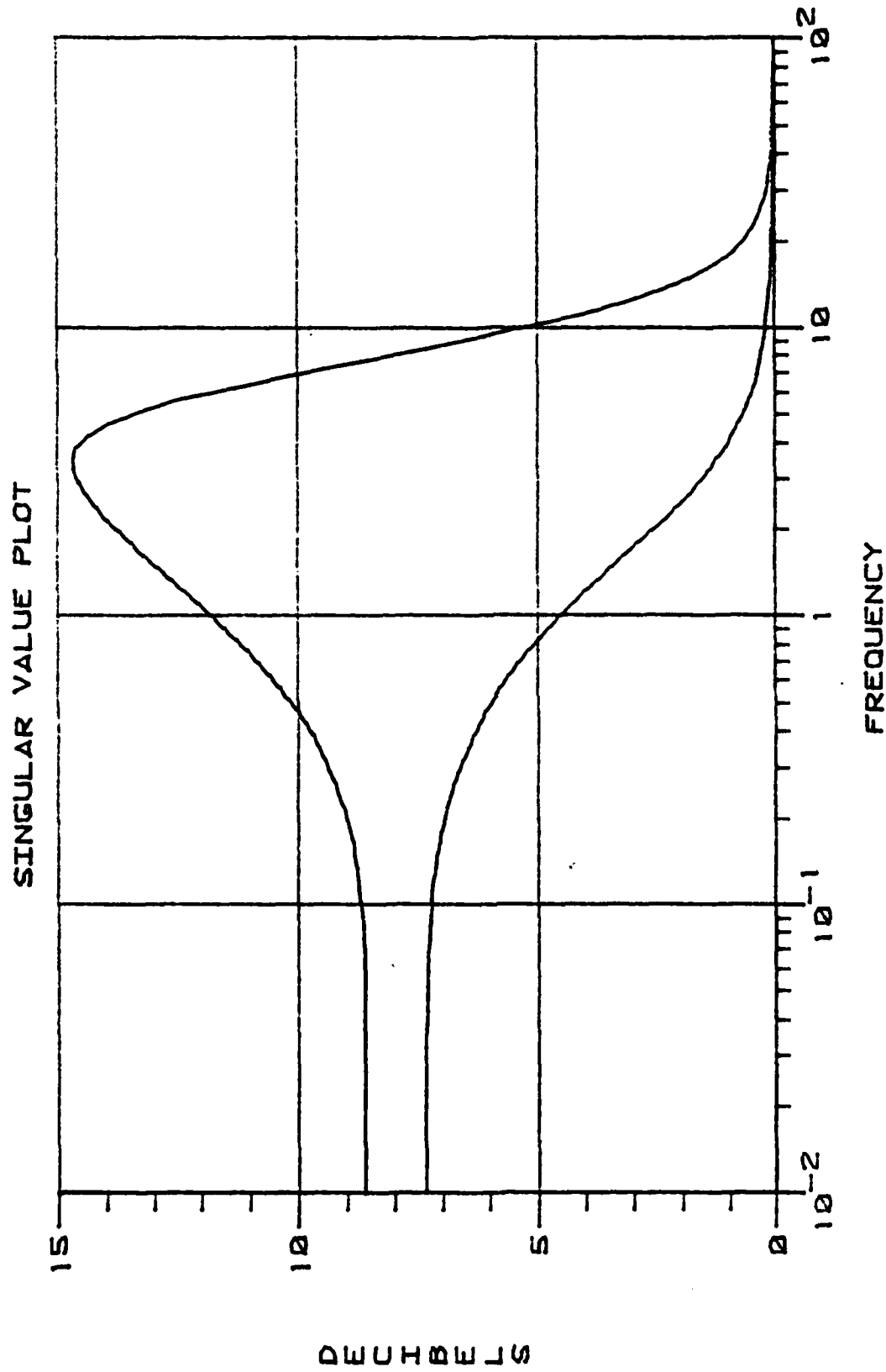
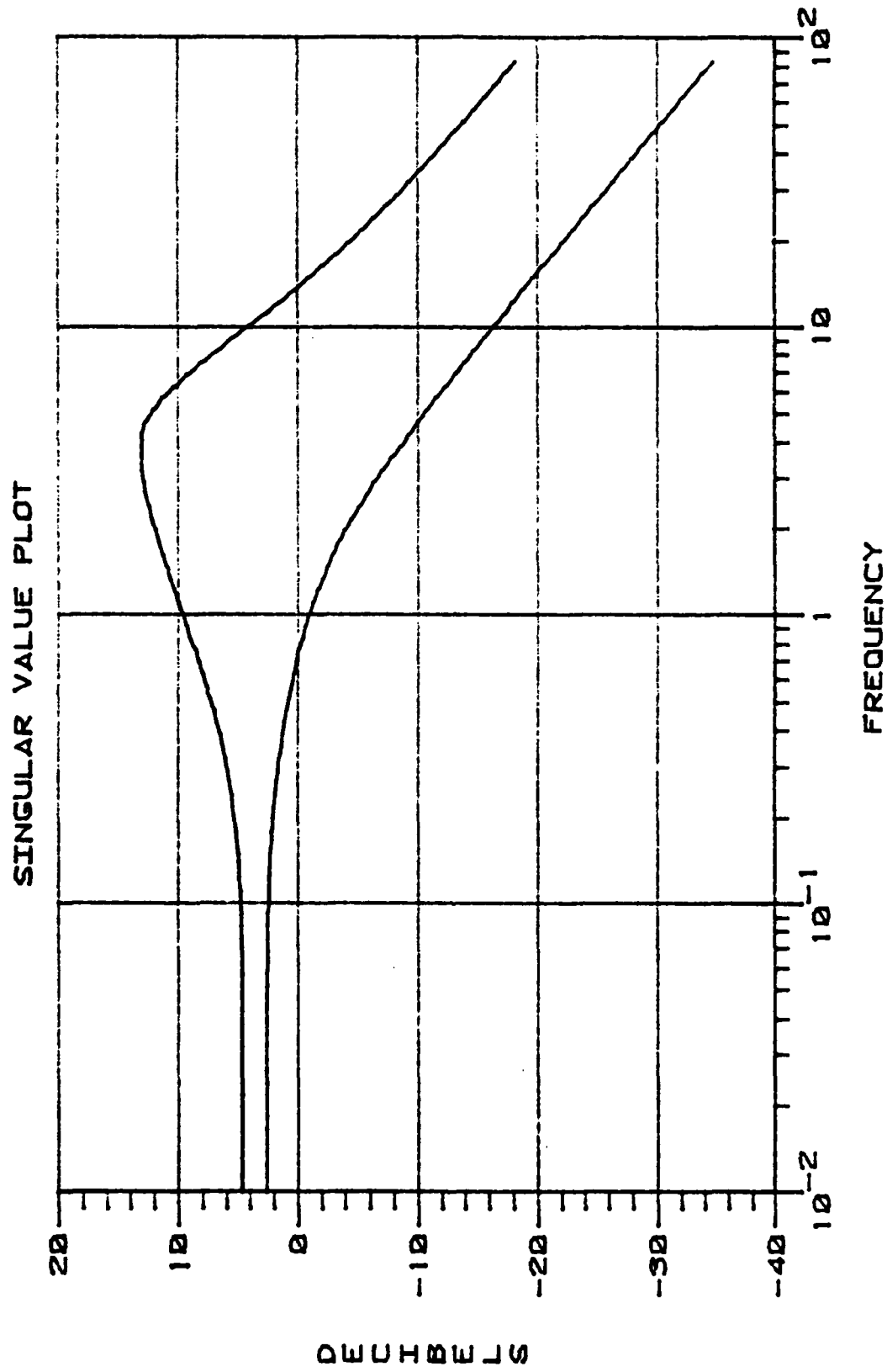Figure 4.8. Singular values of the return difference matrix for the system designed by the gradient method.

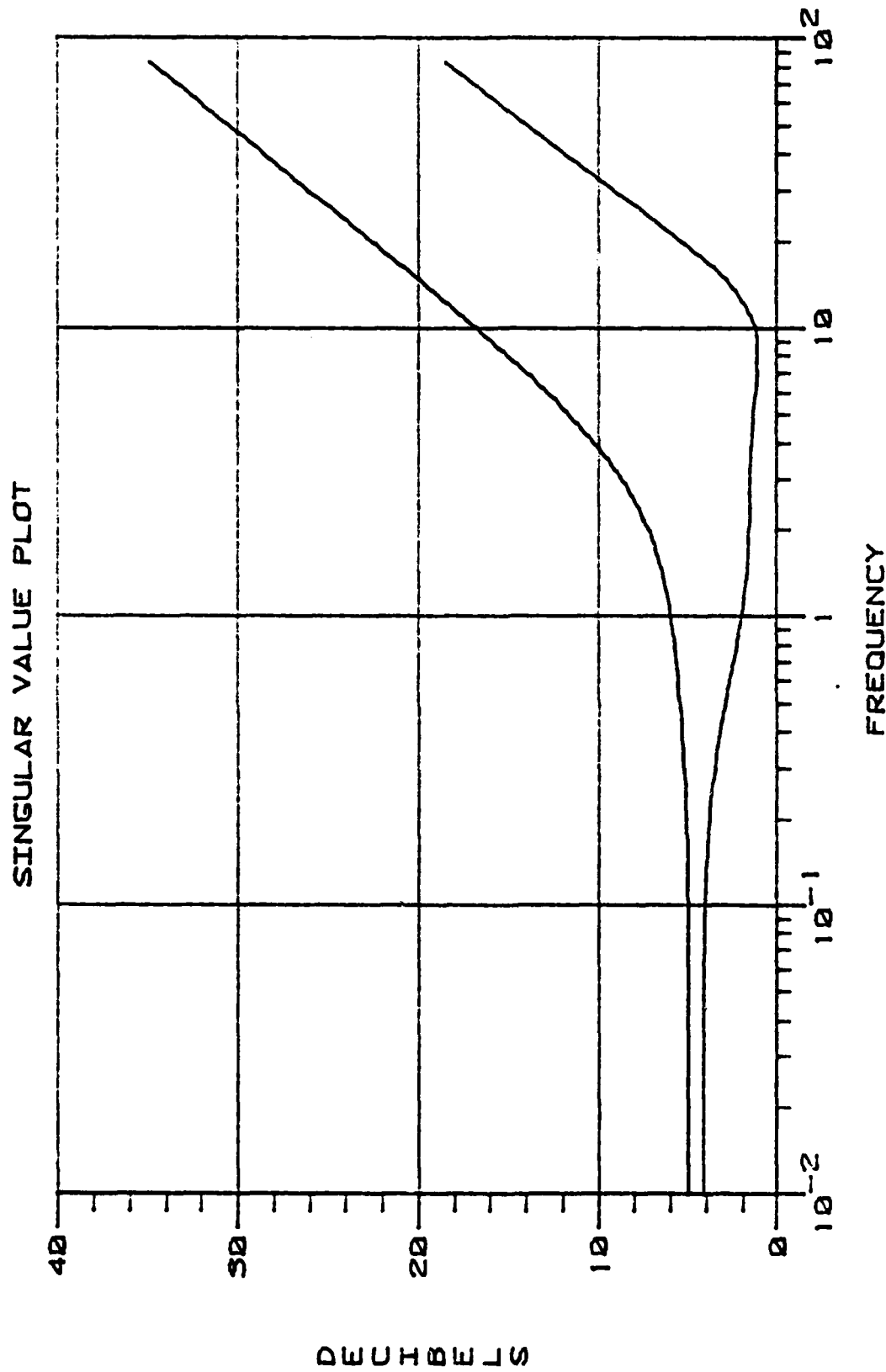Figure 4.9. Singular values of the loop transfer matrix for the system designed by the gradient method.

SINGULAR VALUE PLOT

Figure 4.10. Singular values of the inverse Nyquist matrix for the system designed by the gradient method.

$$|G_2| = 8.31. \tag{4.32}$$

As before, the low frequency value of $\sigma_{A,I,2}$ is 4dB greater than $\sigma_{A,I,0}$. The closed loop bandwidth has increased slightly from the nominal value of $\omega_B \cong 21$ to $\omega_B \cong 22$ rad/sec. The crossover frequency has remained approximately the same as the nominal value at $\omega_c \cong 11$ rad/sec. This indicates that the tolerance of the gradient designed system to high frequency multiplicative errors has remained about the same as the nominal system.

Comparing the eigenvalue placements of the three designs, we see that the asymptotic system has become faster than the nominal system in that the eigenvalues at $-8.01 \pm j8.24$ have moved out to $-11.3 \pm j11.3$. The gradient designed system shows that the extra control energy needed to push the eigenvalues to the left is unnecessary, since the same value of $\sigma_{A,I}$ was achieved in the gradient design where the fast eigenvalues moved only slightly to the left. That less control energy is needed in the gradient designed system is reflected in the fact that the magnitude of the gain matrix G_2 in equation (4.32) has increased by only 17% from the nominal value in equation (4.27).

The size of the G matrix is also important in relative robustness and sensitivity measurements, where the norm of the largest allowable perturbation is compared to the norm of the nominal plant from Figures 4.1 and 4.2

$$\frac{\|P_A\|}{\|P^*\|} = \frac{\underline{\sigma}(I + G_o \varphi B)}{\bar{\sigma}(G_o \varphi B)} = \frac{\sigma_{A,I,0}}{\bar{\sigma}_{L,I,0}} \cong .83, \quad \omega < .1. \tag{4.33}$$

Hence, the nominal system could be subjected to a low frequency ($\omega<.1$) additive perturbation with a magnitude 83% as large as the nominal system

itself, and remain stable. Also,

$$\frac{\sigma_{A,I,0}^{-1}}{\bar{\sigma}_{L,I,0}} \doteq .76 \qquad \omega < .1 \qquad (4.34)$$

and so the nominal system would be insensitive to perturbations with up to 76% magnitude variation. In the asymptotically designed system,

$$\frac{\sigma_{A,I,1}}{\bar{\sigma}_{L,I.1}} \doteq .59 \quad , \qquad \frac{\sigma_{A,I,1}^{-1}}{\bar{\sigma}_{L,I,1}} \doteq .34, \qquad \omega < .1. \qquad (4.35)$$

For this last design,

$$\frac{\sigma_{A,I,2}}{\bar{\sigma}_{L,I,2}} \doteq 1.33 \quad , \qquad \frac{\sigma_{A,I,2}^{-1}}{\bar{\sigma}_{L,I,2}} \doteq .76 \qquad \omega < .1. \qquad (4.36)$$

Notice that in each case above, the magnitude of the nominal plant $(\bar{\sigma}_{L,I,i})$ is different because the G_i matrix is part of it. Because of this, the asymptotic technique actually decreased the relative robustness and sensitivity tolerances. Since $\|G_2\|$ and hence $\bar{\sigma}_{L,I,2}$ did not increase much in the last design, the relative robustness margin was increased from the nominal value.

The fact that $\|G_2\|$ did not increase much over $\|G_0\|$ is also reflected by the low frequency value of $\bar{\sigma}_{A,I,2}$, which is approximately the same as $\bar{\sigma}_{A,I,0}$. This indicates an important feature of the design technique of Section 3 - the motion of the singular values can be separated. If it was necessary to increase the value of $\bar{\sigma}_{A,I,0}$, it could also be done with this technique. Hence the designer has the option of manipulating each system singular value individually.

5. CONCLUSION

In Section 3 we saw how important closed loop properties such as robustness and sensitivity translate into requirements on the singular values of the return difference and inverse Nyquist matrices. This in turn imposes requirements on the singular values of the loop transfer matrix. Thus a major objective in design is to synthesize a controller that adjusts the frequency shapes of the singular values of the loop transfer matrix to meet the requirements imposed by the design criteria. If the mathematical structure of the controller is predetermined, it is possible to calculate the gradient of the singular value functional with respect to the adjustable parameters of the controller. The gradient vector gives the direction to change the control parameter vector (whose elements are the adjustable parameters) that will have the greatest possible effect on a given singular value. Since the control parameter vector is assumed to be constant, a fixed frequency must be substituted into the gradient formula. Then, we are guaranteed only that the singular value functional will be affected in some neighborhood of that frequency. To adjust the entire singular value curve as a function of frequency, the gradient must be iteratively calculated at different frequencies, and the control parameter vector adjusted each time.

The linear quadratic optimal control problem can be used as a method to obtain a full state feedback controller for a general design problem. This method has the advantage of being a systematic, numerically feasible way to choose a feedback gain matrix that has certain minimum guaranteed robustness and sensitivity margins. However, the design problem is now translated into how to choose the weighting coefficients (the Q,R

matrices). The dependence of the system singular values on these matrices
is certainly nonobvious. Since the mathematical structure of the controller
is known, the gradient of these singular value functionals with respect to
the Q,R matrix pair can be calculated. This provides the designer with a
tool to adjust the Q,R matrix pair to bend the singular values (as functions
of frequency) into a more desirable shape.

The new parameter vector (or Q,R pair, in this case) is formed by
adding the gradient times a scalar to the old parameter vector. In the
optimal gradient method, the value of the scalar is determined by considering
the cost functional as a function of the scalar and setting its derivative
equal to zero. The complexity of the relationship between the singular values
and the scalar parameter prevent us from doing this in the LQ problem. Thus
the best value of the scalar must be determined by trial and error in this
design procedure.

In Theorem 3.4, it was necessary to assume that the matrix B had
rank n to insure the existence of the inverse of $T_0 = QBG$ (i.e. the system
had at least as many inputs as it did states). This requirement is not
likely to be satisfied in practice. It may be possible to relax this
assumption by using the pseudoinverse. This possibility is currently under
investigation.

The LQ design presumes that full state feedback is available.
In practice, this is often not the case, and the missing states must be
created by an observer. Unfortunately, when an observer is inserted in
the LQ loop, all robustness and sensitivity guarantees are lost. This
makes the issue of how to adjust the LQ feedback matrix to increase robust-
ness and sensitivity properties much more important. The calculation of

the gradient with an observer in the loop is a straightforward extension of the formulas for the full state feedback case. However, the effectiveness of these formulas remains to be experimentally verified.

When white, uncorrelated Gaussian noise signals are inserted into the system and state measurement equations of the LQ regulator, the optimal solution is the linear-quadratic-Gaussian (LQG) controller. The controller is just the LQ feedback matrix plus an observer whose gains depend upon a dual Riccati equation [23]. The dual equation is a function of the noise covariance matrices. It is possible to calculate the gradient of the singular value functionals witn respect to these matrices using the same type of derivation as in Section 3. This could be useful it the dominant LQ poles are fixed by some primary design constraint. The effect of altering the noise covariance matrices would be that the observer would no longer provide optimal noise reduction. Hence, the robustness and sensitivity properties of the LQG system could be affected at the expense of noise performance.

In Section 4 the LQ adjustment procedure was applied to an example to illustrate its effectiveness. In this example, the additive perturbation was assumed to surround the plant and the controller. This assumption was made because it makes the singular value bound simpler. If the perturbation added directly to the nominal plant, we would have had to take the gradient of the singular value ratio

$$\frac{\underline{\sigma}(I + G\ B)}{\bar{\sigma}(G)} \tag{5.1}$$

(see equation (2.19)). This would complicate the gradient calculation quite

a bit. Also, this bound (5.1) is conservative, which makes it less
meaningful. However, the example showed that it was possible to increase
$\underline{\sigma}(a,i,o)$ for $\omega < .1$ while keeping $\bar{\sigma}(G)$ relatively constant just using the
gradient of $\underline{\sigma}$ (i,o) since this gradient decoupled the motion of the singular
values. Thus, a more complicated gradient calculation of equation (5.1)
could be avoided. Note that this problem does not occur in the case of
multiplicative perturbations – the frequency plot of $\underline{\sigma}(m,\cdot,o)$ provides a
direct nonconservative measure of the desired properties.

REFERENCES

1. H. W. Bode, <u>Network Analysis and Feedback Amplifier Design</u>, Van Nostrand, New York, 1945.

2. I. M. Horowitz, <u>Synthesis of Feedback Systems</u>, Academic Press, New York, 1963.

3. M. G. Safonov, A. J. Laub, and G. L. Hartmann, "Feedback Properties of Multivariable Systems: The Role and Use of the Return Difference Matrix," <u>IEEE Trans. on Automatic Control</u>, Vol. AC-26, 1981, pp. 47-65.

4. J. C. Doyle, "Robustness of Multiloop Linear Feedback Systems," <u>Proc. 17th IEEE Conf. on Decision and Control</u>, San Diego, Calif., January 1979, pp. 12-17.

5. J. B. Cruz, Jr. and W. R. Perkins, "A New Approach to the Sensitivity Problem in Multivariable Feedback Design," <u>IEEE Trans. on Automatic Control</u>, Vol. AC-9, July 1964, pp. 216-223.

6. I. Postlewaite, J. M. Edmunds, and A. G. J. MacFarlane, "Principal Gains and Principal Phases in the Analysis of Feedback Systems," <u>IEEE Trans. on Automatic Control</u>, Vol. AC-26, February 1981, pp. 32-46.

7. J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," <u>IEEE Trans. on Automatic Control</u>, Vol. AC-26, February 1981, pp. 4-16.

8. J. B. Cruz, Jr., J. S. Freudenberg, and D. P. Looze, "A Relationship Between Sensitivity and Stability of Multivariable Feedback Systems," <u>IEEE Trans. on Automatic Control</u>, Vol. AC-26, February 1981, pp. 66-74.

9. M. Athans and P. L. Falb, <u>Optimal Control</u>, McGraw-Hill, New York, 1966.

10. C. A. Harvey and G. Stein, "Quadratic Weights for Asymptotic Regulator Properties," <u>IEEE Trans. on Automatic Control</u>, Vol. AC-23, 1968, pp. 378-387.

11. H. Kwakernaak and R. Sivan, <u>Linear Optimal Control Systems</u>, Wiley-Interscience, 1972.

12. N. A. Lehtomaki, N. R. Sandell, and M. Athans, "Robustness Results in Linear-Quadratic Gaussian Based Multivariable Control Designs,"

13. G. W. Stewart, <u>Introduction to Matrix Computations</u>, Academic Press, New York, 1973.

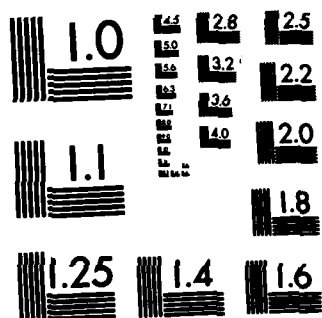14. A. G. J. MacFarlane and D. F. A. Scott-Jones, "Vector Gain," <u>Int. J. Control</u>, Vol. 29, No. 1, 1979, pp. 65-91.

END

FILMED

DTIC

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

15. N. R. Sandell, Jr., "Robust Stability of Multivariable Feedback Systems," _Proc. 16th Annual Allerton Conf. on Communication, Control, and Computing_, Univ. of Illinois, Urbana, October 1978, pp. 471-479.

16. D. G. Luenberger, _Optimization by Vector Space Methods_, Wiley, New York, 1969.

17. C. T. Chen, _Introduction to Linear Systems Theory_, Holt, Rinehart, and Winston, New York, 1970.

18. J. H. Wilkinson, _The Algebraic Eigenvalue Problem_, Clarendon, Oxford, 1965.

19. J. S. Freudenberg, D. P. Looze, and J. B. Cruz, Jr., "Robustness Analysis Using Differential Sensitivities," _Proc. 20th IEEE Conf. on Decision and Control_, San Diego, Calif., December 1981.

20. B. D. O. Anderson, "The Inverse Problem of Optimal Control," Stanford Electronics Laboratories Technical Report SEL-66-038 (T.R. No. 6560-3), Stanford, Calif., April 1966.

21. J. C. Doyle, "Guaranteed Margins for LQG Regulators," _IEEE Trans. on Automatic Control_, Vol. AC-23, August 1978, pp. 756-757.

22. G. Stein and J. C. Doyle, "Singular Values and Feedback: Design Examples," _Proc. 16th Annual Allerton Conf. on Communication, Control, and Computing_, Univ. of Illinois, Urbana, October 1978, pp. 461-478.

23. K. J. Astrom, _Introduction to Stochastic Control Theory_, Academic Press, New York, 1970.

END

FILMED

3-83

DTIC